Medium security units

Jeandarme, Inge

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

MEDIUM SECURITY UNITS

RECIDIVISM & RISK ASSESSMENT

ISBN 978-94-6299-481-2

Printing Ridderprint BV

MEDIUM SECURITY UNITS

RECIDIVISM & RISK ASSESSMENT

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
Tilburg University
op gezag van de rector magnificus,
prof. dr. E. H. L. Aarts,
in het openbaar te verdedigen ten overstaan van een
door het college voor promoties aangewezen commissie
in de aula van de Universiteit

op woensdag 30 november 2016 om 14.00 uur

door

Ingeborg Sophie Joseph Ghislaine Jeandarme,
geboren op 23 februari 1963 te Sint-Truiden, België

Promotores

Prof. dr. S. Bogaerts

Prof. dr. M.S. Groenhuijsen

Prof. dr. T.I. Oei

Overige leden van de promotiecommissie

Prof. dr. G. Meynen

Prof. dr. H. Nijman

Prof. dr. C.L. Mulder

Prof. dr. T. Pham

Dr. M. Spreen

The research presented in this dissertation was performed at Tilburg Law school, Tilburg university, the Netherlands, and at the Knowledge Center Forensic Psychiatric Care, Rekem, Belgium.

Voor Sam en Max

Paranimfen

Hilde Vannoppen

Gokhan Goktas

TABLE OF CONTENTS

Outline

This dissertation consists of ten studies regarding internees and the internment measure in Flanders, Belgium. Internment is a safety measure for offenders who are found not guilty by reason of insanity (NGRI). As in other countries (Every-Palmer et al., 2014; Gordon & Lindqvist, 2007; Salize & Dressing, 2007), this legislation allows offenders with a mental disorder to be transferred to (forensic) psychiatric facilities. The first chapter describes aims, research questions, as well as a careful overview of the various chapters presented.

AIMS OF DISSERTATION

The internment measure has a dual function: the protection of society and the treatment of a forensic special need population. Ultimately, the main goal of forensic psychiatric treatment is to reduce recidivism. Therefore, in many ways, forensic psychiatry differs from general psychiatry. Forensic psychiatric units focus on criminogenic needs and treat reluctant, unmotivated patients. Violence is not uncommon in these patients, and specific skills are required in the staff. Surprisingly, in Flanders, forensic psychiatric units were only established in 2001, 70 years after the Internment Act was in place. This was considered a cornerstone in the development of a forensic psychiatric care system. After years of criticism on the internment system and several convictions by the European Court of Human Rights (ECtHR)[1] for keeping internees in prison without appropriate psychiatric treatment, there has been growing interest from politicians, clinicians, along with public opinion. Since 1995, innovative changes were proposed on two tracks: new proposals for law reform were made, and the care system for internees was improved both inside and outside the prison system, the ultimate goal being a transfer of all NGRI offenders from the prison system.

My work entails research on the first ten years after the establishment of forensic psychiatric units in Flanders. In essence, the whole Flemish population treated within a medium security unit (98%) has been studied here. This population forms an important group, first because of their complex risk and need profile, and second because treatment may be most effective in this group (Skeem, Manchak, & Peterson, 2011). Despite the importance of systematic data registration for the formulation of a policy, Belgium (like many countries) has a lack of systematic and structured collected statistics on its NGRI offenders (Casselman, 2011; De

Vuysere, 2004; Vandevelde et al., 2011). My overall aim is to contribute to the needs of medium security treatment of NGRI patients, and provide medium security and future forensic projects with corroborative findings. This will inform the development of evidence-based policies and practices.

Specifically, this paper has two interrelated aims that address the medium security population. The first is to provide base rate information on recidivism and risk factors involved during and after treatment. The second is to examine how risk is assessed. This dissertation is the first to systematically examine these questions in a Flemish population. Both perspectives may offer valuable information for clinicians and policy makers on how to further conceptualize and implement categorical forensic care.

RESEARCH QUESTIONS

This dissertation focuses on the entire medium security population, and in some chapters on subpopulations. The different studies provide a detailed picture of this newly-used forensic medium security treatment. A series of empirical studies were conducted. Quantitative research methods and literature reviews were used to answer the following research questions:

1. What is the internment measure?
2. How does the evaluation of criminal responsibility in Belgium differ from that of other countries?
3. Who are the medium security internees?
4. Who are the medium security internees initially convicted and later on administratively interned?
5. Which incidents occur during medium security treatment and what is the judicial response?
6. What are the risk factors for inpatient violence?
7. Do internees with psychopathy differ from other internees during medium security treatment?
8. What is the field validity of the Historical, Clinical, Risk management-20 (HCR-20)?
9. What is the field validity of the Psychopathy Checklist-Revised (PCL-R)?

10. What are the reconviction and revocation rates after treatment?
11. What are the risk factors for violent recidivism after treatment?
12. Who are the victims of internees?

OUTLINE OF DISSERTATION

The chapters are clustered into six sections, each corresponding to one or several of the research questions described above. First, I start with legislative background information (Part I, Chapters 1 & 2) and a description of the medium security population (Part II, Chapters 3 & 4). Part III (Chapters 5, 6, & 7) specifically focuses on inpatient incidents during medium security treatment, whereas part IV (Chapters 8 & 9) deals with the reliability and predictive validity of risk assessment instruments during and after medium security treatment. Part V (Chapters 10 & 11) reports on recidivism data and risk factors associated with it after medium security treatment. As stated, the dissertation ends in Part VI with a note on victims during and after medium security treatment (Chapter 12).

Part I. Legislative background

Chapter 1 discusses the internment act and the functioning of the review board (Commission of the Protection of Society, CPS) responsible for the implementation of the internment measure. This background information is important, since medium security units are obliged to work within the scope of the applicable law and (release) decision process of the CPSs.

Chapter 2 focuses on one of the ongoing criticisms of the internment system, i.e., the poor assessment quality of criminal responsibility. A law reform on the internment measure is currently in progress, as there is debate about necessary criteria for both experts and expert reports. So in this chapter, a literature review examines differences between countries concerning legal frameworks and procedures for conducting assessments. Positive practices in other countries are discussed.

Part II. Medium security population

Chapter 3 addresses the study population and provides an overview of socio-demographical, clinical, judicial, and criminogenic features of the medium security population. Elements of a theoretical model regarding the difference between a low, medium, and high security population are discussed.

Chapter 4 specifically focuses on a subpopulation of the medium security population, namely convicted internees, or convicted detainees who are later administratively interned by the Minister of Justice when a serious mental illness emerges during their detention. A previous study in Wallonia identified problems with this population, which was assessed as more dangerous. In this chapter, we compare psychiatric disorders and risk profiles of convicted internees and regular internees in Flanders - to investigate whether similar results apply to the Flemish population. The upcoming law reform to abolish the administrative internment of convicted detainees has been heavily criticized on legal grounds. Clinical implications of the abolishment are also discussed.

Part III. Inpatient medium security treatment

Chapter 5 examines incidents reported to the CPS during medium security treatment, including the subsequent judicial reaction to these incidents. Studying nonviolent and (verbally) violent incidents is important, since each can affect the treatment process and result in early treatment termination. Forensic psychiatric patients have traditionally been stigmatized as more violent, more difficult to treat, and less compliant than other patients. The establishment of medium security units raises concern for safety in the community, so this chapter examines if these presumptions can be confirmed. In addition, it looks at whether incidents coded under offending categories yield new convictions.

Chapter 6 presents the results of research on risk factors associated with one particular form of critical incident, namely verbal and physical violence during medium security treatment. Violence is a frequent phenomenon in psychiatric and forensic psychiatric settings, and can have a profound impact on staff and patients. The underexposure of inpatient violence and victimization in the literature is surprising, because one of the primary objectives of forensic psychiatric treatment is to teach violent patients how to de-escalate, and help vulnerable patients, such as

psychotic forensic patients, prevent (re)victimization. Furthermore, inpatient violence predicts violent recidivism after treatment. Retrospectively, nonviolent patients are compared to patients who engaged in (verbally and physically) violent behavior during medium security treatment. Multivariate analyses are used to assess which variables independently contribute to a higher risk of violence.

Chapter 7 describes the treatment course of a subpopulation with strong psychopathic traits. According to the Risk Need Responsivity model, intensive care and supervision is required for patients with psychopathic traits. However, there is much reluctance to take these people into treatment, not only because of the supposed limited chances of success, but also because therapy-interfering behavior is expected. Despite the fact that medium security units excluded internees with psychopathic traits at first, 75 internees with psychopathy were hospitalized and treated during the period from 2001 to 2010. In this chapter, differences between internees with and without high PCL-R scores are examined and discussed in line with international findings.

Part IV. Risk assessment during and after medium security treatment

In **Chapter 8,** one of the primary tasks of forensic treatment - the assessment of risk for further violent offending – is scrutinized. Structured risk assessment is part of routine practice in forensic settings. However, little attention has been paid to the clinical applicability of existing risk assessment tools. This study focuses on the performance of the HCR-20 in practice. The predictive validity of violent recidivism, during and after medium security treatment, is evaluated.

Chapter 9 investigates the interrater reliability (IRR) and predictive validity of the PCL-R. Previous studies found evidence of reduced IRR and predictive validity when the PCL-R was used in applied settings. Thus, despite tremendous research demonstrating good IRR and predictive validity in research settings, the reliability and validity of the PCL-R in the field is still a major point of debate. The number of field validity studies is limited, and reproducibility of results is even more important, since the PCL-R can have a major impact on the judicial process and treatment options. Scoring discrepancies between hospital and prison settings, as well as differences in predictive validity across these two settings, are also examined.

Part V. Recidivism after medium security treatment

Chapter 10 examines the criminal outcome after medium security treatment and is the first study reporting on recidivism rates of internees in Flanders. Treatment outcome in forensic mental health is best measured over a broad range of areas, including clinical and humanitarian issues. However, the prevention of future criminal behavior is the most important goal in forensic psychiatric treatment, so studying recidivism rates is critical. In this chapter, reconviction rates are considered a reliable measure of recidivism. We also conduct an analysis on which internees were granted unconditional release by the CPSs. Conditional release relating to recidivism of these two subgroups (internees under conditional release and those who received unconditional release) are appraised.

Chapter 11 further explores a second recidivism outcome measure in a subpopulation, namely medium security internees under the authority of CPS Ghent. Along with official reconviction data, incidents falling under offending categories are reported to the CPS, and added in order to provide a more detailed picture of relapse. In addition, risk factors for this combined outcome measure are presented.

Part VI. Victims

Chapter 12 relates to the ultimate goal of all forensic treatment, namely, the prevention of further victimization. This study examines victims of the index internment measure as described in Chapter 3. Knowledge on victim characteristics (age, gender, and relationship to the offender) creates a more effective treatment of internees and the prevention of even more victims. Surprisingly, in contrast to the scholarly attention paid to forensic perpetrators, studies examining victim characteristics are scarce. The limited literature suggests that victims are primarily adults, known to the offender. While previous research focused primarily on descriptive characteristics, this study also examines whether diagnostic categories in the offender moderates these findings.

The thesis ends with a **general discussion**, where main results are summarized and discussed and suggestions for future research are addressed.

FOOTNOTES

[1] In 1998 (Aerts v. Belgium ECHR 1998-V 1939), the ECtHR first censured Belgium for the unlawful detention of internees. Since then, many similar ECtHR rulings followed, including two recent pronouncements censoring Belgium for inhumane handling of vulnerable internees kept in prison for many years, while deprived of adequate psychiatric care: Claes v. Belgium App No. 43418/09 (ECtHR January 10, 2013); Lankester v. Belgium App No. 22283/10, (ECtHR January 9, 2014) (Hanoulle & Verbruggen, 2016).

REFERENCES

Casselman, J. (2011). Internering in Vlaanderen anno 2010: Ondanks het goede nieuws nog steeds mensonwaardig: Liber amicorum Johan Goethals. In I. Aertsen, T. Daems, A. Haaren, F. Hutsebaut, & J. Maesschalck (Eds.), *Deviante wetenschap. Het domein van de criminologie* (pp. 237-247). Leuven: Acco.

Every-Palmer, S., Brink, J., Chern, T. P., Choi, W. K., Hern-Yee, J. G., Green, B., . . . Mellsop, G. (2014). Review of psychiatric services to mentally disordered offenders around the Pacific Rim. *Asia-Pacific Psychiatry, 6*(1), 1-17. doi: 10.1111/appy.12109

Gordon, H., & Lindqvist, P. (2007). Forensic psychiatry in Europe. *Psychiatric Bulletin, 31*(11), 421-424. doi: 10.1192/pb.bp.107.014803

Hanoulle, K., & Verbruggen, F. (2016). Neuroscepticism in the courtroom: The limited role of neuroscientific evidence in Belgian criminal proceedings. In S. Moratti & D. Patterson (Eds.), *Legal insanity and the brain: Science, law and European courts* (pp. 41-75). Portland: Hart Publishing.

Salize, H. J., & Dressing, H. (2007). Admission of mentally disordered offenders to specialized forensic care in fifteen European Union member states. *Social Psychiatry and Psychiatric Epidemiology, 42*(4), 336-342. doi: 10.1007/s00127-007-0159-2

Skeem, J., Manchak, S. & Peterson, J. (2011). Correctional policy for offenders with mental illness: Creating a new paradigm for recidivism reduction. *Law and Human Behavior, 35,* 110–126.

Vandevelde, S., Soyez, V., Beken, T. Vander, Smet, S. De, Boers, A., & Broekaert, E. (2011). Mentally ill offenders in prison: The Belgian case. *International Journal of Law and Psychiatry, 34*(1), 71-78. doi: 10.1016/j.ijlp.2010.11.011

Vuysere, S. De (2004). Wetgevende initiatieven inzake internering. In D. Van Daele & I. Van Welzenis (Eds.), *Actuele thema's uit het strafrecht en de criminologie* (pp. 63-75). Leuven: University Press.

Internment act

1

01

INTRODUCTION

Under Belgian law, internment is an indefinite safety measure and not a punishment, and it has a dual purpose, namely the protection of society *and* the medical-psychiatric treatment of the offender (Vandevelde et al., 2011). After the internment measure is imposed, interned persons fall under the jurisdiction of one of the regional Commissions for the Protection of Society (CPS), which are responsible for the execution of this safety measure. In a Belgian population of around 11 million inhabitants, about 300 to 350 offenders are placed under this internment measure annually (Dienst voor Strafrechtelijk Beleid, 2016). Over the years, the number of internees has increased; at the end of 2013, there were about 3,820 internees in Belgium (Deckers et al., 2014). In what follows, an overview is given of the legislative background of the internment measure, starting with a historical sketch.

HISTORY

The Belgian Criminal Code of 1867 is based on the Napoleonic Criminal Code of 1810, inspired by the principles of the Enlightenment. The deterrent and retributive functions of criminal law presuppose that the potential offender reasons and acts freely and can be held accountable for his actions. However, the legislator foresaw a legal excuse for offenders who commit a crime while being insane. According to Article 71 of the Criminal Code, an offender who is insane or who was compelled to act by an irresistible force (acting under *force majeur* or duress) cannot be found guilty: "Er is geen misdrijf, wanneer de beschuldigde of de beklaagde op het ogenblik van het feit in staat van krankzinnigheid was of wanneer hij gedwongen werd door een macht die hij niet heeft kunnen weerstaan". Because criminal accountability is excluded in Article 71, the law assumes *a contrario* that accountability (and guilt) is a necessary component of the crime. *Mens rea*, in this case guilt, is a necessary pre-condition for being held accountable *(nullum crimen, sine culpa*; no crime without guilt*)*. In cases of *insanity defense*[1] or irresponsibility, there is no guilt in the offender, who is regarded as unable to act under the influence of reason and free will (De Clerck, 1988). Offenders found not guilty due to insanity (in French *démence*) were subsequently referred to so-called *hospices d'aliénés* or *maisons de santé* (institutions for the insane), where they

(aliénés délinquants) were subjected to the same rules as the regular insane.[2] Unlike the offenders found completely irresponsible for their acts, offenders with diminished responsibility *(délinquants aliénés)* were sentenced to prison but served less time in prison because of mitigating circumstances. In addition, offenders with diminished responsibility were sometimes acquitted due to a liberal interpretation of Article 71.

At the end of the 19th century, the way the judicial system dealt with both types of offenders with a mental disorder (délinquants aliénés and aliénés délinquants) generated criticism. It was assessed that these measures were unable to adequately protect society from these offenders and that the hospitals were not equipped for this particular group of patients (Goethals, 1997). Under the influence of a strong *Social Defense* movement that had emerged in Europe, a paradigm shift took place from guilt and (ir)responsibility to social dangerousness and from punishment to *security* (Cosyns, 2001). Criminal justice theories were largely influenced by these social defense ideas. Social defense scholars argued that high recidivism rates proved the failure of classical approaches to criminal justice. Human beings were considered to be far from possessing the rationality and free will that the classical approach attributed to them. Instead, new scientific (e.g., psychological and criminological) research on criminality provided a foundation for a new approach, one that would take into account the characteristics of individual offenders and would be better equipped to protect society. Under the influence of these findings, criminal law underwent significant changes and a *modern movement* of criminal law emerged, especially in western European countries such as Belgium and the Netherlands (Groenhuijsen & van der Landen, 1990). Under the old criminal law, the focus was on the criminal act or *actus reus*. Criminals were punished for what they did, not for who they were. Important principles in this movement were: 1) criminals are individuals with a free will and choose to commit crimes, 2) criminals are guilty of choosing to act in a criminal way and can be held responsible, and 3) criminals deserve to be punished. The primary objectives were retaliation (”an eye for an eye”) and general crime prevention. Regardless of whether the underlying theory was absolute (retaliation, “an eye for an eye”) or relative (general prevention), it was paramount that punishment and crime would match, and therefore proportionality was considered a very important element. An individual was seen as a *homo economicus*, weighting benefits and costs against *a priori* clearly formulated criminal consequences. Punishment and crime were thus

closely matched. Legal thinking was largely influenced by the basic ideas formulated by von Feuerbach: "Nulla poena sine lege, nulla poena sine crimine, nullem crimen sine poena legali" (Groenhuijsen & van der Landen, 1990). In the modern movement it was stated that punishment should serve other purposes and not only retaliation and prevention. An individual was no longer considered to be a person with a free will, choosing to commit crimes. Other elements, either within the individual or outside of them, could influence the individual and induce him to commit a criminal act. Some authors such as Lombroso focused on concepts of physiognomy and internal motives (*born criminal*), whereas other scholars such as Lacassagne emphasized that external factors or contextual pressures precipitated the individual to commit crimes. The modern movement no longer saw the individual as a rational person, but adapted a more deterministic point of view. In other words, it was believed that individuals not always chose to commit a crime. Two important implications of this movement were that: 1) individuals can be forced to commit crimes or do not possess a free will and therefore cannot be held responsible for their actions, and 2) criminals should be helped or treated (i.e., resocialization should be the primary aim of punishment). The focus here is on the perpetrator, not on the act. Other consequences are that the judge is allowed more subjective sanctioning, depending on the individual case (Groenhuijsen & van der Landen, 1990).

Under the influence of one of the leading figures of the modern or social defense movement, Prins, several new laws were introduced in Belgium, such as the Law of Conditional Release in 1888. Prins argued that the available tools were not adequate for some types of offenders, such as recidivists who do not appear to respond to punishment "because of their insensitivity, impulsiveness and lack of reason" and insane people (Hanoulle & Verbruggen, 2016). The modern movement also influenced several initiatives that were implemented within the prison system, such as the establishment of the Penitentiary Anthropological Service and several psychiatric prison wings (*annexes*). Also, during this era in the beginning of the 20th century, several law proposals were made for criminally insane offenders, who were considered not (fully) responsible for their acts but simultaneously dangerous because of their proclivity to crime (Hanoulle & Verbruggen, 2016). The need to adopt a new approach for insane offenders and offenders with diminished responsibility led, after long debates and numerous changes to the original proposal, to the passing of the Act of April 9, 1930 to Protect Society from Abnormals

and Habitual offenders (APS; *Wet van 9 april 1930 tot Bescherming van de Maatschappij tegen Abnormalen en Gewoontemisdadigers, Belgisch Staatsblad 11 mei 1930*). From then on, offenders found not criminally responsible *and* offenders found to possess diminished responsibility were no longer sentenced, but were instead subjected to a new measure, the *internment measure*. The Internment Act was very innovative for its time. For example, not only the completely insane but also the severely disturbed could be subjected to this measure. A medical observation was proposed in order to advice the judge, and CPSs were established for the execution phase of the internment. In 1964, the Act of April 9, 1930 was adapted and replaced by the Act of July 1, 1964 to Protect Society from Abnormals and Habitual Offenders (*Wet van 1 juli 1964 tot Bescherming van de Maatschappij tegen Abnormalen en Gewoontemisdadigers, Belgisch Staatsblad 17 juli 1964*). According to some critics, this law reform was no more than a bare refinement of the old law (Goethals, 1997). However, some important changes were made: the mandatory presence of a lawyer, the possibility to transfer patients to private hospitals outside of the prison system, and the duration of the measure which was changed to an indefinite one (Casselman, 2015).

Internationally, the Belgian law on social defense has been seen as a good law, with the exception of the internal and external legal status of the internees, which was assessed as insufficient (Research voor Beleid, 1995). In recent years, a number of minor changes have been made to the law: mandatory referral to a specialized service in cases of sexual abuse of a minor (Act of April 13, 1995) and the possibility to appeal a rejected request for release (Act of February 10, 1998). Also, in the aftermath of the affair Dutroux, the title of the law was changed to the Act to Protect Society from Abnormals, Habitual Offenders and Offenders who committed particular types of sex offenses (*Wet tot Bescherming van de Maatschappij tegen de Abnormalen, Gewoontemisdadigers en Plegers van bepaalde seksueel strafbare feiten, Belgisch Staatsblad 2 april 1998*) (Goethals & Robert, 2007). Meanwhile, a Commission Internment (1996-1999) was set in place to examine criticisms regarding the internment law and its implementation. The vice-president of the commission, Vandemeulebroeke, prepared a draft in April, 2001, introducing a new law. It took another six years before the new law was accepted. Since April 21, 2007, a new Act on Internment of Persons with a Mental Disorder (*BS* July 13, 2007) amended and replaced the Act of July 1, 1964. The new act, which ultimately was never implemented, was determined to be a highly procedural law, in which the emphasis was more on risk and safety than on treatment

and reintegration (Smets, Verelst, & Vandenberghe, 2009). On May 5, 2014, the internment measure was reformed once again, with the Act on the Internment of Persons with a Mental Disorder (*BS* July 9, 2014). The most recent Internment Act combines some of the changes proposed in 2007, such as the replacement of the CPSs with the Chambers for the Protection of Society (part of the tribunal for the execution of sentences), although it did take into account some of the criticisms raised in the field. Currently, the most recent Internment Act (2014) is – again – undergoing reparations and will be implemented shortly. Until that time, the Act of July 1, 1964 to Protect Society from Abnormals and Habitual Offenders[3] and Offenders who Committed Particular Types of Sex Offenses (hereinafter Act Protection Society or APS) is still in effect.

CURRENT ACT ON SOCIAL DEFENSE

Articles 1 and 7 of the APS lay down the conditions for internment. Unlike the insanity defense, which is based on the mental state of the offender when committing the crime (see supra), the internment is based on the mental state and degree of dangerousness at the time of the judicial decision (Hanoulle & Verbruggen, 2016).

Internment is a safety measure of indefinite duration that can be imposed by the trial courts (i.e., the police tribunal, court for misdemeanors, and assize court) and the investigating courts (i.e., the chambers supervising judicial investigations), the latter with the exception of political and press offenses.

In order to impose an internment measure, several conditions have to be fulfilled. First, the offender must have committed a felony or misdemeanor for which the criminal law sets a minimum penalty of at least eight days (Art. 7 APS). Second, the person must be found "hetzij in staat van krankzinnigheid (state of insanity), hetzij in een ernstige staat van geestesstoornis (a serious state of mental disturbance) of van zwakzinnigheid (mental deficiency) die hem ongeschikt maakt tot het controleren van zijn daden" (Art. 1 APS). In other words, the offender must be found unfit to control his actions due to his mental disturbance and is thus considered not guilty by reason of insanity (NGRI). The act does not formally distinguish between complete irresponsibility or severely diminished responsibility. Third, an important aspect emerging in the jurisprudence (case law of the Court of Cassation[4]) relates to the social dangerousness of the

offender at the time of the assessment. Internment is solely imposed on offenders found socially dangerous at the time of sentencing.

Other than the abovementioned terms, there is no specified legal standard of insanity. When analyzing the specific incapacities mentioned in the law and the practical application of the law, it seems that both cognitive and volitional prongs allow for NGRI acquittal. Which psychiatric disorders qualify for the incapacities is not mentioned explicitly in the law. The legal terms should be interpreted according to their meaning in ordinary language and not the specific meaning they may have in particular technical domains (including the medical domain). According to its ordinary meaning, insanity refers to a mental disorder that causes a complete lack of control over one's actions and undermines one's intellectual capacities (Hanoulle & Verbruggen, 2016). In case law, medical approaches prevail. According to psychiatrists, *zwakzinnigheid* (mental deficiency) refers to moderate or severe intellectual disability, as well as milder forms of intellectual disability with poor social functioning. *Krankzinnigheid* (insanity) is an old and ambiguously defined term, which usually refers to major mental disorders such as psychotic disorders, major depression, bipolar disorder, or dementia. The term *ernstige staat van geestesstoornis* (severe mental disorder) is a broad term, which encompasses a heterogenic group of disorders such as personality disorders, substance use disorders, and paraphilic disorders. Which disorders can be considered a severe mental disorder is up for some debate, although this category is used in more than half of the reports (De Vuysere, Casselman, & Vervaeke, 2004).

Phase 1 internment: Criminal investigation

There are three phases in the internment process: criminal investigation, sentencing, and execution. During a criminal investigation, an investigating judge or public prosecutor can appoint a psychiatrist to assess whether the defendant acted under the influence of one of the states mentioned in Art. 1 APS. Likewise, at trial, the judge can also make such an appointment. While criminal responsibility will be assessed by a psychiatrist in most cases, this is not mandatory, contrary to other European countries (see further Chapter 2). In Belgium, the judge is allowed to hear experts, and there are no legal requirements regarding the scientific or professional training and expertise of these experts, other than that they must be a licensed physician (Hanoulle & Verbruggen, 2016).

A psychiatric report investigating criminal responsibility is usually performed on an ambulatory basis and is in some cases supplemented by psycho-diagnostic examination by a psychologist. An exploratory study by De Clercq and Vander Laenen (2013) found that six out of ten psychiatrists never included psychological testing in their reports, whereas in 67% of the investigated reports, diagnostic instruments (36% intelligence tests and neuropsychological tests, 33% personality assessments, and 31% risk assessment schemes) were used. In theory, the act provides the opportunity for a clinical forensic psychiatric observation, which can take place in a psychiatric prison ward (Art. 1 APS). However, due to a shortage of (trained) personnel, the observation is hardly ever used. The decision by the Belgian government to build a specifically designed clinical observation center – the so-called Penitentiary Research and Clinical Observation Center *(Koninklijk Besluit van 19 april 1999 tot instelling van een Penitentiair Onderzoeks- en Klinisch Observatiecentrum (POKO) met het statuut van Wetenschappelijke inrichting van de Staat, Belgisch Staatsblad 8 mei 1999)* – was not implemented (Vander Beken & Van Steenbrugge, 2010).

Phase 2 internment: Sentencing stage

When it is clear from the available evidence that the suspect has committed the offense, it is up to the judge to decide whether the necessary conditions for an internment measure are met. If the offender is found to be not criminally responsible at the time of the sentencing, an internment measure can be imposed, regardless of the degree of (in)sanity at the time of the crime. If the offender is found not responsible at the time of the offense but responsible at the time of the sentencing, the offender cannot be found guilty of the perpetrated crime. In those cases, there will be an acquittal on the basis of Article 71 SW, whereas the 1964 Act forms the basis for the internment measure. When the conditions for internment are met, according to the Court of Cassation, the judge does not have to assess whether the insanity defense applies at the time of the offense, but in practice this will usually be done in order to obtain a wider understanding of the mental state and dangerousness of the offender (Hanoulle & Verbruggen, 2016).

Internment is not a punishment, nor can it be combined with a criminal penalty, but appealing the decision is possible. Unlike many other countries (e.g., the Dutch Entrustment Act),

Belgium does not have a system of diminished responsibility with possibilities for combining protective measures with criminal sanctions.

The decision of internment is not exclusively made by the trial judge but also at the level of the investigating courts[5], who will then act as a trial court (Casselman, 1997). Between 1965 and 1980, 73% of internment measures were imposed by an investigating court and only 27% by a trial court (Goethals, 1985). A more recent study found that 68% of internment measures were imposed by an investigating court, 24.3% by a trial court, and 7.8% by the Minister of Justice (De Vuysere et al., 2004).

In addition to the regular internment measure, the Minister of Justice can order a ministerial or administrative internment measure when a convicted prisoner develops a severe mental disorder in the course of detention (Art. 21 APS). The decision by the Minister of Justice is made with the unanimous consent of the CPS and is not open for appeal. The *Article 21 procedure* has been the subject of heavy legal criticism (see further Chapter 4). In summary, the current possibilities for the imposition of an internment measure are presented in Figure 1.

In the 60s, internment measures were imposed quite regularly (500 times a year), whereas the number of new internments dropped to 300 internments per year in the 70s and has since remained constant, between 300 and 350 annually. Relative to the total number of sentences for

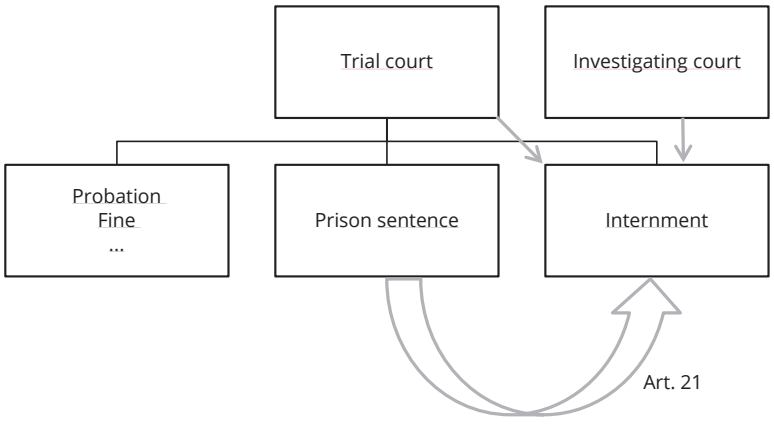


Figure 1. Pathways towards an internment measure.

similar offenses the percentage of internment measures remains small, namely 1 or 2% (Delva, 1999; Goethals, 1985). This represents approximately 0.2 to 0.3% of all convictions (including, for example, traffic offenses; Dienst voor Strafrechtelijk Beleid, 2016; Ministerie van Justitie, 2014). However, the total number of internees has been rising each year. In 2004 there were 3,306 internees in Belgium and in 2009 the numbers of internees increased to 3,956 (Vander Beken & Van Steenbrugge, 2010). Moens and Pauwelyn (2012) counted 4,093 Belgian internees in February, 2011, with faster growth in Flanders than in Wallonia. The most recent numbers showed a small decline in the total number of internees, with about 3,820 internees at the end of 2013 (Deckers et al., 2014).

Phase 3 internment: Execution

After the internment measure is imposed, internees fall under the jurisdiction of one of the regional CPSs. The CPS is a dedicated body that functions as a court of law. It is chaired by a judge and has two additional members, a lawyer and a psychiatrist. The public prosecutor is present at the hearings but not at the deliberation. A clerk provides administrative help. There are three Flemish CPSs (CPS Antwerp covering the judicial districts Antwerp and Limburg, CPS Ghent covering the judicial districts West and East Flanders, and CPS Leuven covering the judicial district Leuven), four French CPSs (Jamioulx, Lantin[6], Mons, Namur), one mixed CPS (Vorst), and one *high CPS* for appeal (Brussels). Each month, the high CPS addresses 10 to 15 appeals, which are usually filed by personality disordered internees from CPS Antwerp (Vandenbroucke, 2009). In 2004, CPS Antwerp was the largest Flemish CPS, responsible for more than half of all Flemish internees (Casselman, Devuysere, & Vervaeke, 2003a; Cosyns, 2005).

The CPS is primarily responsible for the execution of the internment measure and has extensive powers. Some of the important decisions to be made by the CPS include:

- appointing the setting in which the internee will be treated (Art. 14 APS);
- granting temporary leave (Art. 15 APS);
- granting conditional and unconditional release (Art. 18 APS);
- determining the duration of conditional release (Art. 20 APS); and
- advising the Minister of Justice on the administrative internment (Art. 21 APS).

After the internment decision, the internee appears for the first time before the CPS. At the first hearing, the CPS decides the setting in which the internee will be treated. Mandatory treatment can be provided either within the prison system or outside of the prison system. Inside of the prison system, the internee is placed by the CPS (Art. 14 APS) in a psychiatric ward or Institution for the Protection of Society. Outside of the prison system, the CPS can refer the internee to a private hospital, either under the internment measure (Art. 14 APS or *placement*) or under the conditional release from internment measure (Art. 18 APS). On September 15, 2004, 60.5% (*n* = 1,999) of internees (*N* = 3,306) had been conditionally released, and 39.5% (*n* = 1,307) were interned (Cosyns, 2005). Differences between Flemish CPSs were noted, although these differences were not statistically examined (Casselman et al., 2003a): CPS Antwerp had more incarcerated internees (33.2% compared to 28.2% for Ghent and 25.0% for CPS Leuven) but also more internees in ambulatory care (35.4% compared to 29.1% in Ghent and 34.9% in Leuven), whereas CPS Leuven (34.2%) and Ghent (38.4%) had more internees in private institutions than CPS Antwerp (23.4%).

Automatic hospitalization is not required after the internment decision, since conditional release into the community is also an option. Figure 2 summarizes the placement possibilities under the internment measure (Art. 14 APS) and the different referral options for conditional internment (Art. 18 APS). Thus, according to the specific treatment needs (low, medium, or high care), risk of recidivism (low, medium, or high), and security level (low, medium, or high), as assessed by the judicial officer or the psychosocial prison team, internees can – at least in theory – either reside in prison, forensic psychiatric units, regular psychiatric units, protected houses, or they can reside in the community while receiving ambulatory care. Before the start of the first *medium security units* in Flanders in 2001, forensic beds were not available. It was only recently that the first *high security* forensic hospital was established (FPC Ghent, since the end of 2014). As a consequence, many Flemish internees deemed too dangerous for community supervision or low to medium security care still remained in prison without adequate treatment (Deckers et al., 2014; Moens & Pauwelyn, 2012; Vandevelde et al., 2011). The number of internees remaining in prison has grown in the past decades, from 589 internees in March, 1999 to 1,087 in December, 2013, as shown in Figure 3 (Federale Overheidsdienst Justitie, 2011, 2012; Macquet, 2014; Moens & Pauwelyn, 2012). In Belgian prisons, internees are the third largest category of detainees, and

their number is growing more rapidly than other groups of inmates. In 1999, 7.5% of the total prison population were internees, and in 2011 this rose to 10% (12% in Flanders and 8% in Wallonia; Federale Overheidsdienst Justitie, 2011; Moens & Pauwelyn, 2012).

Upon conditional release, the patient's liberty is dependent on the adhering to several requirements, which usually include living in a particular treatment setting and continuing to receive psychiatric supervision and treatment (*sociaal-geneeskundige voogdij*). As soon as the security level has diminished, the internee will either be referred to a lower security (inpatient) facility or released into the community with outpatient care. All treatment options are coupled with a judicial mandate to receive treatment and follow all specified rules. Should a breach of

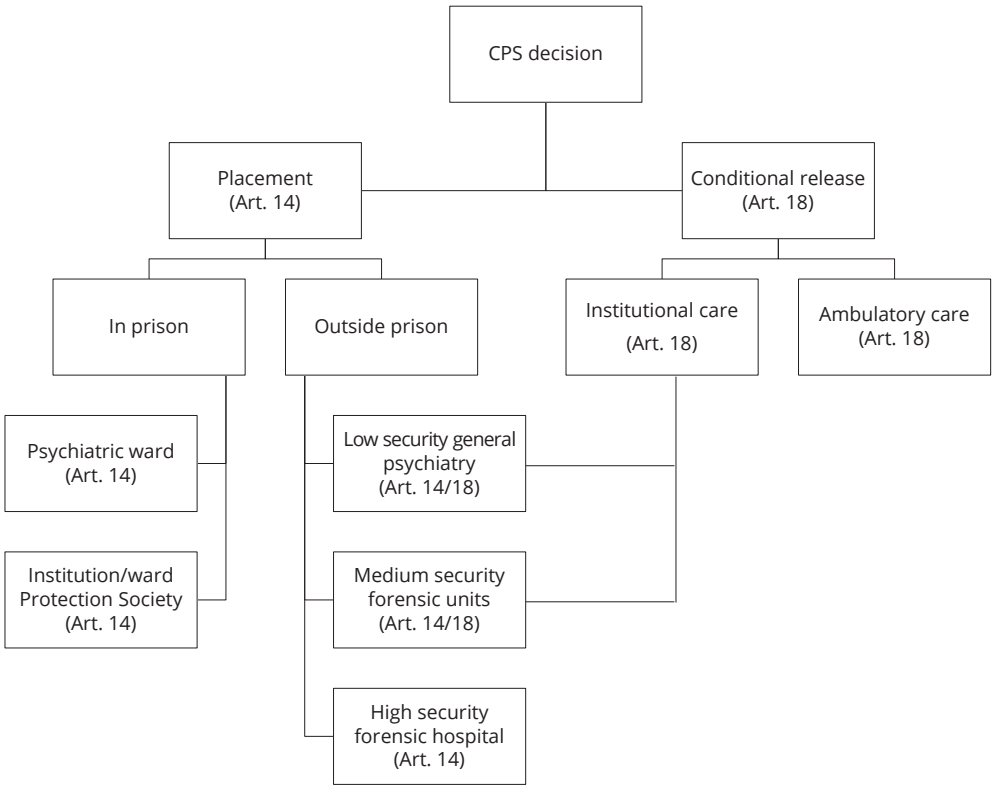


Figure 2. Execution of internment measure.

conditions occur, the internee can be readmitted to prison on the request of the public prosecutor, and the conditional release will be revoked (Art. 20 APS). As soon as the internee is no longer a risk for society, the CPS will release the internee unconditionally and stop the internment measure. After that, the CPS will have no authority over them anymore. A forensic psychiatric examination prior to unconditional release is not mandatory. The CPS can terminate the internment measure indefinitely if the mental condition of the internee has sufficiently improved and the conditions are met for their rehabilitation. In most cases, unconditional release is preceded by a period of conditional release, of which the duration is not legally determined but is set by the CPS (usually between three and five years). The conditional release period can be extended several times if necessary. According to Belgian law, the duration of internment is undefined and remains enforceable until the offender's mental health problems are resolved (Vandevelde et al., 2011).

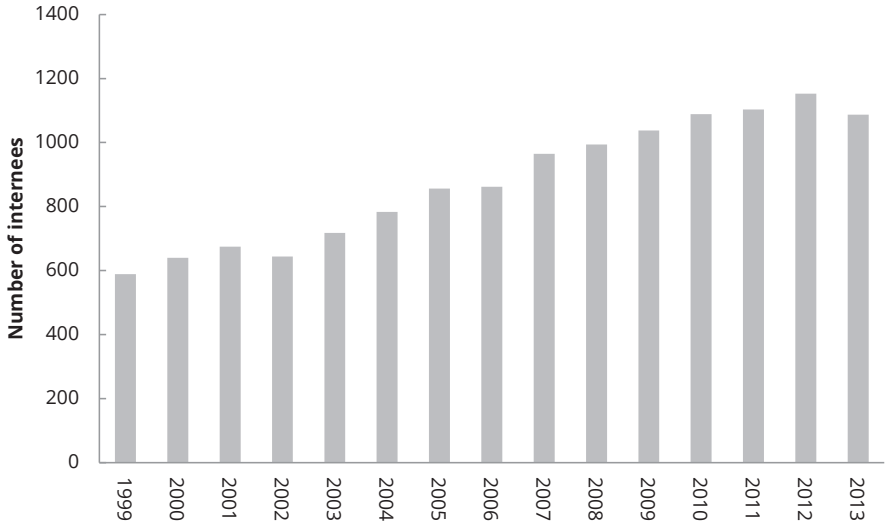


Figure 3. Number of interned inmates (situation on March 1, 1999–2011, on August 7, 2012 and on December 23, 2013).
Sources: Federale Overheidsdienst Justitie (2011, 2012); Macquet (2014); Moens and Pauwelyn (2012).

INTERNMENT IN THE SOUTHERN PART OF BELGIUM

National data concerning internees are very scarce, but at least one study, conducted under the authority of Cosyns, provided a one-day prevalence snapshot. On September 15, 2004, the situation of the entire interned population differed between the southern and the northern part of the country (Cosyns, 2005). In Flanders, 70.4% of the internees (*n* = 1,203) were conditionally released as compared to 49.9% (*n* = 796) in Wallonia (see Figure 4).

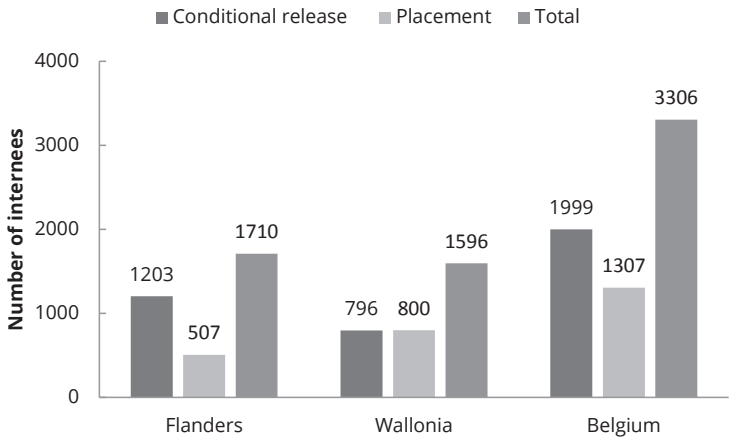


Figure 4. Number of internees September 15, 2004 (Cosyns, 2005).

Figure 5 shows the situation of internees without conditional release in greater detail. Internees without conditional release can be divided into incarcerated internees (who receive little to no treatment) and internees placed in institutions that provide psychiatric treatment. In Flanders, almost all (99.6%, *n* = 505) of the internees who were not conditionally released were incarcerated in services pertaining to the Federal Public Service of Justice (i.e., psychiatric or regular wards of a prison, as well as institutions and departments of Social Defense), with no or very limited therapeutic interventions. In Wallonia, this was the case for less than half (42.9%, *n* = 343) of the internees. In Flanders, only two internees (0.4%) were placed in private hospitals outside of the prison system, whereas 83 internees (10.4%) in Wallonia were placed in private psychiatric institutions or sheltered homes and 374 internees (46.8%) resided in Facilities for

Social Defense (*Les Maronniers*, Tournai and *Le Chêne aux Haies*, Mons), governed by the Walloon region, with psychiatric care comparable to general psychiatric units.

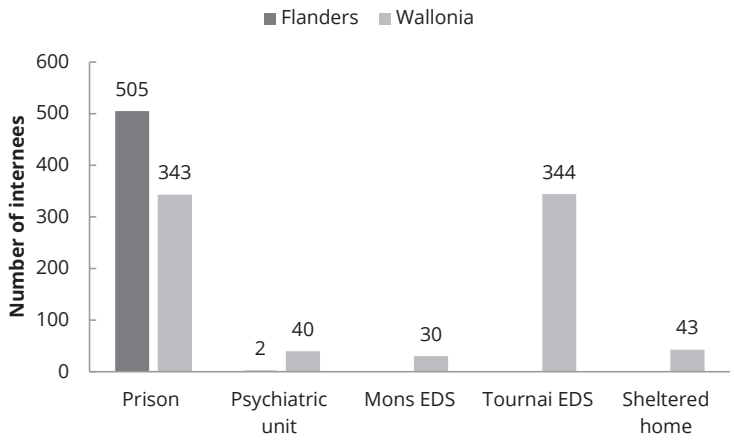


Figure 5. Place of residence internees not conditionally released on September 15, 2004 (Cosyns, 2005).

Figure 6 shows the situation of internees with conditional release more in detail. Again, substantial differences between Flanders and Wallonia can be noted. The majority of conditionally released internees in Wallonia remained in the community (96.1%, *n* = 675), while a substantial proportion of conditionally released internees in Flanders (38.7%, *n* = 466) resided within an institution (Cosyns, 2005).

Taken together, these data clearly reflect a different policy and care system in the two parts of the country (Cosyns, D'Hont, Janssens, Maes, & Verellen, 2007). The Flemish CPSs established a working relationship with different private organizations, enabling the transfer of internees under conditional release (Art. 18 APS) without providing extra resources. The French CPSs referred the majority of their internees to two psychiatric institutions for Social Defense (Tournai and Mons), which were subsidized by the Minister of Justice (Art. 14 APS). In addition, another *Etablissement de Défense Sociale* was opened by the Minister of Justice in Paifve in 1972, which can be regarded as the first prison-hospital institution for French-speaking internees (Casselman,

2015). Because this dissertation focuses exclusively on Flemish internees, the situation in Wallonia will not be further explored (for more information please refer to Cartuyvels, Champetier, & Wyvekens, 2010).

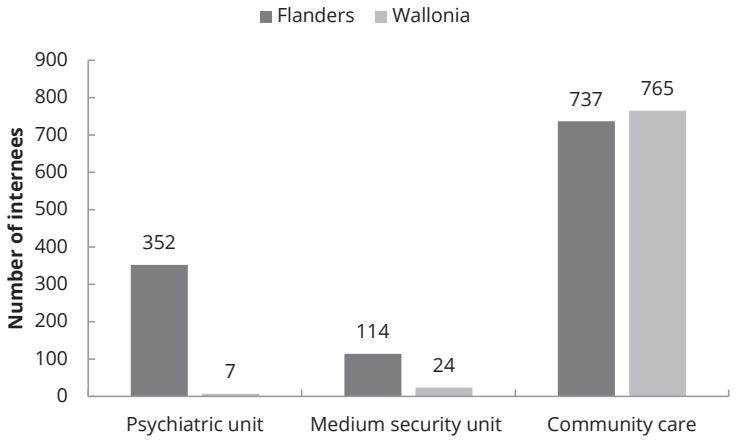


Figure 6. Place of residence internees conditionally released on September 15, 2004 (Cosyns, 2005).

CRITICISM ON THE CURRENT LAW ON SOCIAL DEFENSE

Several criticisms have been raised regarding the internment measure. First, the poor quality and lack of criteria on both experts and expert reports have been noted by several authors (De Clercq & Vander Laenen, 2013; Vandevelde et al., 2011). Other problems relating to criminal responsibility evaluations include the lack of multidisciplinary and clinical observation and low financial compensation, which are often cited as reasons for the shortage of experts. When collecting the data from the CPS judicial files in the current study, the quality of the expert reports varied substantially. But even more surprisingly, very little information was available at every step of the execution of the internment measure.

The second criticism relates to the dichotomized model in which offenders are declared either criminally responsible or irresponsible. This excludes the possibility of combining protective measures for criminal penalties (Vandevelde et al., 2011). Also, the distinction between

normality and psychopathology is not clear-cut, but rather a continuous measure from mild to severe mental disability (Casselman, 2003). The information on judicial antecedents and recidivism rates found in one of the current studies showed that even for the same offenders, both types of sentences are pronounced quite regularly, suggesting some subjective interpretation occurs. Also, in practice, internment results in incarceration *and* treatment periods, however without due process, which again suggests rather subjective or pragmatic decisions.

Third, the lack of a uniform policy for deciding upon internment has been raised. According to Goethals (1985), this resulted in different internment rates across jurisdictions. The lack of a common policy across different CPSs was also noted at the execution phase by several authors (Casselman, Devuysere, & Vervaeke, 2003b; De Ruyver & Goethals, 1991; Delva, 1999; Goethals, 1985). For example, Goethals (1985) found that prior to 1980, CPS Antwerp, Ghent, and Vorst had the highest numbers of conditional releases to the community at first court appearance, whereas CPS Leuven had the highest number of referrals to private institutions. Another more recent study found that 50% of the internees in CPS Leuven were conditionally released at first appearance compared to 26% in CPS Antwerp, 17% in CPS Ghent, and 3.5% in CPS Mons (Delva, 1999). The Commission Internment was very critical regarding diversity and the stereotyped character of the different CPS decisions. Furthermore, the commission argued that the composition of the CPS was not well balanced and that there was a lack of training and continuing education (Delva, 1999). Vandenbroucke (1981) questioned the expertise of the members of the CPS and their conservative release policy, as internees were often revoked for minor violations. Despite these criticisms, it was generally recognized that the caseload of CPSs was too high and that there was a lack of resources, and in general personnel operated under difficult work circumstances. After collaborating with the different CPS administrations when collecting data for the current studies, the vast majority of the CPSs were found to be very cooperative and open-minded toward the research. Also, although this was not the topic of the current research, differences between CPSs were examined (most analyses not shown) and revealed practically no differences between the CPSs. In other words, the current study could not confirm this assumption.

Fourth, the internal and external legal status of the internees was found to be inadequate. According to Vandenbroucke (1981) internment often leads to a very lengthy to lifelong measure,

with few possibilities for appeal. Also, more specifically, the administrative internment procedure was heavily criticized on its legal grounds. The current study shows that this criticism is confirmed, for example when assessing revocation rates (see Chapter 10).

The main criticism of the current internment law, however, has been the lack of treatment opportunities, both inside and outside the prison system (Broothaerts & Mentens, 1982; Casselman, 1997; Cosyns et al., 2007; De Vuysere, 2005; Delva, 1999; Goethals, 1997). The European Committee for the prevention of Torture (CPT) and the International Observatory of Prisons (OIP) both held the Belgian government in violation of international standards for its serious systemic shortcomings regarding the situation of internees in prison. Units inside of the prison system were found to be outdated and clearly insufficient and the chance of successful reintegration into the community was considered very small, due to the shortage of forensic psychiatric beds (Naudts et al., 2005). Since 1998, Belgium has been convicted 22 times by the European Court of Human Rights (Casselman, 2015). It is not entirely clear why it took so long for the government to establish forensic psychiatric beds in Flanders. Possible explanations relate to the lack of funding based on poor political courage, tensions between the judicial and care discourse, and the multiple stigmatization of the population at hand. The establishment of the first categorical forensic medium security units in 2001 and the first high security hospital in 2014 were seen as a big step forward. However, they can hardly be regarded as novel ideas, since in 1873 the building of a *prison-hospice* had already been proposed, which would be a special institution with both prison and psychiatric hospital features (Casselman, 2015). The question also remains why the development of the system was strikingly different in both parts of the country, also possibly reflecting less political courage in Flanders. However, while the organization was less problematic in Wallonia, Flanders appears to have caught up quite quickly in the last decade.

Finally, several authors noted a striking lack of systematically planned and implemented research at all stages of the internment process (Cartuyvels et al., 2010; Casselman, 1997; Cosyns, 2006; Decoene, 2010; Vandenbroucke, 1994). Even the most basic statistics are not available. In 2011, Vandevelde et al. (2011) noted that the most recent figures could be found in the 2007 annual report of CPS Ghent, which found that one out of three internees falling under the jurisdiction of CPS Ghent were incarcerated, while the number of internees residing in private

hospitals was growing. The literature reviews conducted in the course of this dissertation revealed little and fragmented research, and no information on criminal or other outcomes measures, which corroborates this last criticism. Furthermore, the current research was financially subsidized by the Public Psychiatric Care Center Rekem (*Openbaar Psychiatrisch Zorgcentrum)* and *Limburg Sterk Merk*, because to date governmental funding has not been available.

CONCLUSION

Internationally, the Belgian law on social defense was evaluated as a good law, with the exception of the internal and external legal status of the internees, which was assessed as insufficient (Research voor Beleid, 1995). Nationally, the criticism has primarily focused on the poor quality of the psychiatric reports and the lack of clinical observation. Assessing criminal responsibility is not an easy task (see Chapter 2), but nevertheless a very important one. It is the entrance gate to an indefinite protection measure.

Right from the beginning, the internment measure had a double aim. First, society had to be protected from offenders with a mental disorder, and second, it was evident that offenders should receive adequate treatment in order to ensure a safe transition to the community afterwards. However, the law was not always applied properly, which has led to a growing number of internees in prison. Several law reforms, published in 2007 and 2014, have not yet been implemented and were therefore not discussed. Furthermore, law reforms provide no guarantees for the solution of the most fundamental problem of the internment measure, namely the inadequate care system (Goethals & Robert, 2007).

Since 2000, the Belgian government has taken several initiatives – especially in Flanders – to expand treatment options for interned offenders. Amongst others, forensic medium security units were established and many other initiatives have emerged (see Chapter 3). The medium security units, which are the subject of the current dissertation, were very important because they were the start of the categorical forensic care system.

FOOTNOTES

[1] It should be noted that in the Belgian inquisitorial system, insanity is not a defense on the procedural level, but rather an exclusion for the culpability of the defendant on the substantive level (Hanoulle & Verbruggen, 2016).

[2] From 1850 on, the law on civil commitment was used for the mandatory treatment of insane people.

[3] For certain categories of offenders, such as recidivists and habitual offenders, a *Terbeschikkingstelling van de Regering (TBR)* could be added to a prison sentence at the sentencing stage. In practice, after having served their main sentence, these persons were mostly still considered as inmates, residing in ordinary prions. Besides imposing additional conditions upon release, the Minister of Justice could also decide to impose an internment, which is not the same as the internment measure. It is not a measure of indefinite duration and the CPSs are not responsible for the application. In 2006, the TBR was removed from the Internment Act and changed into *Terbeschikkingstelling door de Strafuitvoeringsrechtbank (TBS)*.

[4] Cass February 26, 1934, *Pas* 1934, I, 180.

[5] Investigating courts (*onderzoeksgerecht*) are chambers that supervise judicial investigations. The investigating court functions as a filter between the investigation phase and the trial phase. It has the authority to refer the suspect to the trial court or dismiss the charges and release the suspect. When it is clear that the suspect has committed the offense, the investigating court can impose internment upon the request of the suspect or of the public prosecutor.

[6] CPS Lantin was closed in 2003 and the CPS files were transferred to CPS Namur (De Vuysere, 2004).

REFERENCES

Beken, T. Vander, & Steenbrugge, W. Van (2010, October). *Internering. Recht voor hulpverleners*. Paper presented at the Just Care symposium, Ghent, Belgium.

Broothaerts, J., & Mentens, R. (1982). Zijn er in de huidige context behandelingsmogelijkheden op een afdeling sociaal verweer? *Panopticon, 3*(5), 435-441.

Cartuyvels, Y., Champetier, B., & Wyvekens, A. (2010). La défense sociale en Belgique, entre soin et sécurité. Une apporche empirique. *Déviance et société, 34*(4), 615-645.

Casselman, J. (1997). Internering: Huidige situatie. In J. Casselman, P. Cosyns, J. Goethals, M. Vandenbroucke, D. De Doncker, & C. Dillen (Eds.), *Internering* (pp. 39-61). Leuven/Apeldoorn: Garant.

Casselman, J. (2003). Forensisch Psychiatrisch deskundigenonderzoek. In W. Van de Voorde, J. Goethals, & M. Nieuwdorp (Eds.), *Multidisciplinair forensisch onderzoek. Juridische en wetenschappelijke aspecten* (pp. 453-473). Brussel: Uitgeverij Politeia.

Casselman, J. (2015). De interneringswetgeving in historisch perspectief. Komt er nooit een einde aan de Sisyfusarbeid? In J. Casselman., R. De Rycke, & H. Heimans (Eds.), *Internering. Nieuwe interneringswet en organisatie van de zorg* (pp. 5-25). Brugge: Die Keure.

Casselman, J., Vuysere, S. De, & Vervaeke, G. (2003a). De Commissies ter Bescherming van de Maatschappij van Antwerpen, Gent en Leuven. Een vergelijking van de geïnterneerdenpopulatie. *Panopticon, 6*, 604-609.

Casselman, J., Vuysere, S. De, & Vervaeke, G. (2003b). De geïnterneerdenpopulatie van Vlaanderen: Enkele cijfergegevens. *Panopticon, 6*, 599-604.

Clerck, D. De (1988). *Beginselen van strafrecht en strafvordering*. Leuven: Acco.

Clercq, M. De, & Laenen, F. Vander (2013). Gebruik van testinstrumenten in psychiatrische deskundigenverslagen bij internering; Een exploratief dossieronderzoek in het gerechtelijk arrondissement Gent [The use of psychodiagnostic tests in forensic psychiatric reports within the framework of internment. An exploratory study of records on internees in the judicial district of Ghent]. *Tijdschrift voor Psychiatrie, 55*(5), 337-347.

Cosyns, P. (2001). De behandeling van geïnterneerden: Een doorbraak in een nieuwe wet? *Verhandelingen, LXIII*, *63*(4), 305-318.

Cosyns, P. (2005). *Werkgroep forensisch psychiatrisch zorgcircuit: Syntheseverslag mei 2005*. Brussel: Federale Overheidsdienst Justitie.

Cosyns, P. (2005, September). *Behandeling van geïnterneerden in Vlaanderen: Een uitdaging voor de toekomst.* Paper presented at the Herfstvergadering VVP, Grimbergen, Belgium.

Cosyns, P., Peteghem, P. Van, Raes, B.C., & Sabbe, B.G. (2006). De forensische psychiatrie in Vlaanderen op een keerpunt? [Forensic psychiatry in Flanders at a crossroad?]. *Tijdschrift voor Psychiatrie, 48*(7), 511-513.

Cosyns, P., D'Hont, C., Janssens, D., Maes, E., & Verellen, R. (2007). Geïnterneerden in België. De cijfers. *Panopticon, 1*, 46-61.

Deckers, A., Seynnaeve, K., Smedt, A. De, Dheedene, J., Vandenplas, E., & Auwera, A. Van der (2014). *Geïnterneerden in detentie op 24/12/2013*. Unpublished report.

Decoene, S. (2010). Lange zwangerschap, uitgestelde geboorte: Over de geschiedenis en toekomst van de forensische psychologie in Vlaanderen. *Tijdschrift Klinische Psychologie, 3*, 169-179.

Delva, J. (1999). *De commissie internering. Eindverslag van de werkzaamheden*. Unpublished government report, Ministry of Justice.

Dienst voor Strafrechtelijk Beleid. (2016). *Statistieken*. Retrieved from http://www.dsb-spc.be/web/index.php?option=com_wrapper&Itemid=129

Federale Overheidsdienst Justitie. (2011). *Justitie in cijfers 2010*. Retrieved from http://justitie.belgium.be/nl/publicaties/justitie_in_cijfers_2010

Federale Overheidsdienst Justitie. (2012). *Justitie in cijfers 2011*. Retrieved from http://justitie.belgium.be/nl/publicaties/justitie_in_cijfers_2011

Goethals, J. (1985). Internering in België. Enkele statische gegevens. *Panopticon*, 129-147.

Goethals, J. (1997). De wet tot bescherming van de maatschappij in een historisch perspectief. In J. Casselman, P. Cosyns, J. Goethals, D. Vandenbroucke, D. De Doncker, & C. Dillen (Eds.), *Internering* (pp. 11-38). Leuven/Apeldoorn: Uitgeverij Garant.

Goethals, J., & Robert, L. (2007). Een nieuwe interneringswet. *Fatik, 114*, 4-9.

Groenhuijsen, M.S. & Landen, D. van der (1990). De I.K.V. in het spanningsveld tussen klassieke rechtsbeginselen en moderne rechtsopvattingen. In M. S. Groenhuijsen & D. van der Landen (Eds.), *De moderne richting in het strafrecht* (pp. 7-93). Anrnhem: Gouda Quint bv.

Hanoulle, K., & Verbruggen, F. (2016). Neuroscepticism in the courtroom: The limited role of neuroscientific evidence in Belgian criminal proceedings. In S. Moratti & D. Patterson (Eds.), *Legal insanity and the Brain: Science, Law and European courts* (pp. 41-75). Portland: Hart Publishing.

Macquet, D. (2014). *Cartographie des usagers 'internés'. Présentation du projet (avant réalisation) et du contexte auprés du comité sectoriel 'santé' du CPVP*. Paper presented at the comité sectoriel 'santé' du CPVP.

Ministerie van Justitie. (2014). *Statistieken van de hoven en rechtbanken 2014*. Retrieved from http://justitie.belgium.be/nl/informatie/statistieken/hoven_en_rechtbanken/2014

Moens, I., & Pauwelyn, L. (2012). *Geen opsluiting, maar sleutels tot re-integratie. Voorstellen voor een gecoördineerd zorgtraject voor geïnterneerden*. Brussel: Zorgnet Vlaanderen.

Naudts, K. H., Cosyns, P., McInerny, T., Audenaert, K., Eynde, F. van den, & Heeringen, C. van (2005). Belgium and its internees: A problem for human rights and a stimulus for service change. *Criminal Behaviour and Mental Health, 15*(3), 148-153. doi: 10.1002/cbm.1

Research voor Beleid, Maastrichts Europees Instituut voor Transnationaal Rechtswetenschappelijk Onderzoek. (1995). *Voorzieningen voor de behandeling en bejegening van psychisch gestoorde delinquenten; een internationaal vergelijkende studie*. Leiden: Research voor Beleid, METRO.

Ruyver, B. De, & Goethals, J. (1991). Actiedag internering, Antwerpen, 5 oktober 1990. *Panopticon*, *12*(2), 184-189.

Smets, H., Verelst, R., & Vandenberghe, J. (2009). Geestesziek en gevaarlijk: gedwongen opname of internering? Het Belgische wettelijke kader [Mentally ill and dangerous: civil commitment or internment? The Belgian judicial framework]. *Tijdschrift voor Psychiatrie, 51*(4), 217-225.

Vandenbroucke, M. (1981). Bemerkingen bij het functioneren van de internering. *Panopticon, 1*, 80-85.

Vandenbroucke, M. (1994). Een eenheid voor forensische psychiatrie in een psychiatrisch ziekenhuis: Utopie en realiteit. *Panopticon, 15*(1), 76-88.

Vandenbroucke, M. (2009). De wereld van de internering en van de geïnterneerden: Het functioneren van de Commissie ter Bescherming van de Maatschappij. *Panopticon, 30*(6), 92-97.

Vandevelde, S., Soyez, V., Beken, T. Vander, Smet, S. De Boers, A., & Broekaert, E. (2011). Mentally ill offenders in prison: The Belgian case. *International Journal of Law and Psychiatry, 34*(1), 71-78. doi: 10.1016/j.ijlp.2010.11.011

Vuysere, S. De (2004). Wetgevende initiatieven inzake internering. In D. Van Daele & I. Van Welzenis (Eds.), *Actuele thema's uit het strafrecht en de criminologie* (pp. 63-75). Leuven: University Press.

Vuysere, S. De (2005). Veertig jaar wet ter bescherming van de maatschappij: Reden tot feesten? In J. Goethals, F. Hutsebaut, & G. Vervaeke (Eds.), *Gerechtelijke geestelijke gezondheidszorg: Wetenschap, beleid en praktijk. Liber amicorum Joris Casselman* (pp. 259-269). Leuven: University Press.

Vuysere, S. De, Casselman, J., & Vervaeke, G. (2004). Geïnterneerden verblijvend in een rijksinstelling in Vlaanderen. *Tijdschrift voor Welzijnswerk, 260*, 18-35.

Criminal responsibility evaluations: Benchmarking in different countries

Submitted

Jeandarme, I. [†]*, Pouls, C.* [†]*, Al-Taiar, H., Brink, J., Canton, W., Kristiansson, M., Thibaut, F., Verreyt, V., & Konrad, N.*

[†] These authors contributed equally to this work.

2

ABSTRACT

Forensic clinicians play a crucial role in criminal responsibility evaluations. However, the quality of these assessments has at time come under scrutiny and has been heavily criticized. A literature review revealed significant differences between countries concerning legal frameworks and procedures for conducting these assessments. The findings suggest that although some countries can be seen as a *role model*, there still is room for improvement.

INTRODUCTION

Forensic psychiatrists, as well as other mental health professionals, provide the legal system with clinical information and assessments concerning the suspect's functioning, mental state and capacities at the time of the alleged offense and/or trial. These forensic assessments play a crucial role in Court, influencing subsequent decision-making on sentencing, placement or treatment of mentally disordered offenders. However, studies that put evaluations on criminal responsibility in an international perspective are scarce, focusing mostly on the adjudication and treatment of forensic psychiatric patients (Bal & Koenraadt, 2000; Dressing & Salize, 2006; Every-Palmer et al., 2014). The role of these so-called expert witnesses has been subject of increasing criticism (Wettstein, 2005), often because of highly contested cases in which conflicting psychiatric assessments are presented. Clear legal frameworks, procedures, formats and appropriate training have therefore been recommended (Dressing & Salize, 2006).

This study compared several, mainly European, countries in order to explore best practices in the assessment of criminal responsibility. The following countries were selected for this matter: Belgium, the Netherlands, France, Germany, England, Sweden and Canada.

METHOD

The information was primarily gathered through (non-) peer reviewed literature, reports and legislation. A literature search was performed in PubMed and PsychINFO using the following keywords: criminal responsibility (reports/evaluations), pre-trial assessment, psychiatric expert, (forensic) psychiatric assessment, court/sentencing, sanity evaluation and insanity defense. Additional articles were identified through reference lists. In addition, experts (forensic psychiatrists) from the respective countries were approached to comment on the findings.

FINDINGS

A literature search revealed substantial differences in the assessment of criminal responsibility between the countries.

Criteria evaluation

The main features of evaluations of criminal responsibility in individuals accused of a crime are presented in Table 1. Not guilty by reason of insanity (NGRI) is the most common legal term used in the contributing countries, while in Canada the term *Not Criminally Responsible on account of Mental Disorder* has been used since 1992 (Every-Palmer et al., 2014). In some countries such as Belgium, an examination of the mental state by an expert is not mandatory in order for the judge to absolve an offender of his responsibility. To reach a verdict of unaccountability or in case compulsory treatment is considered, an expert opinion is mandatory in the Netherlands, Germany, England, Canada, Sweden and in some cases in France (mandatory in case of severe offenses or a sexual crime) (Bal & Koenraadt, 2000; Every-Palmer et al., 2014). Even if not mandatory, it is common practice that a judge will order an assessment to assist him in his decision-making (Dressing & Salize, 2006).

This assessment order can be made at any stage of the proceedings. An exception to this is Sweden, where the concept of unaccountability was abolished in 1965. In Sweden everyone is considered to be responsible for his or her actions, implicating that the presence of a severe mental disorder (and thus also the assessment order) is discussed at a later stage, after the offender has been found guilty or has confessed to the crime and before the sentencing stage (Belfrage & Fransson, 2000). Furthermore, no country applies specific (legally defined) criteria on the basis of when an assessment order should be made. Only in the Netherlands a decision-supportive instrument, called *Beslissingsondersteuning Onderzoek Geestesvermogens* (BooG; van Kordelaar, 2002), was developed to advice the judicial authorities/court about the necessity and type of assessment to be done. At this point, the use of the instrument is optional, but there is reason to believe that it will be mandatory in the near future (NIFP, personal communication).

Most countries provide for the opportunity to keep the defendant under clinical observation in forensic psychiatric assessment units in order to provide a more elaborated multidisciplinary report about his/her mental state. In Sweden these units are governed by the National Board of Forensic Medicine, a governmental agency. Only in the Netherlands, a special observation unit, called *Pieter Baan Center*, was designed for these clinical observations. In Belgium, there is a legal provision to do so, but so far, there are no such institutions.

Only Sweden uses a mandatory format for the report. Other countries have several guidelines that have been developed, describing the *ideal* format (Boetticher, Nedopil, Bosinski, & Saß, 2007; Fédération Française de Psychiatrie, 2007; McLeod, Sweeting, Joyce, Evans, & Barkley, 2010; Nederlands Instituut voor Forensische Psychiatrie en Psychologie, 2014). These guidelines are quite similar and in accordance with the general guidelines pronounced by the American Academy of Psychiatry and the Law (Glancy et al., 2015) and the World Psychiatric Association (2015) like for example, collecting collateral information, separating findings from opinion and the use of standardized test.

Finally, except for Sweden, all psychiatric evaluations assess insanity, which is a legal rather than a medical definition. Legal insanity typically refers to two main features: a) a reference to mental disease or defect and b) an analysis of the specific incapacities – cognitive and/or volitional – allowing for acquittal by reason of insanity. There are few differences in the legal standard for insanity, with most countries including the cognitive component.

Criteria expert

With the exception of Belgium and Germany, all countries stipulate selection criteria with regard to the qualifications, professional training or experience of the expert (Table 2). In the Netherlands and Sweden these criteria are rather strict and mandatory, with an external institution supervising and safeguarding the quality, reliability and competence of the expert.

Although not always legally defined, all of the abovementioned countries consult at least one psychiatrist when criminal responsibility is questioned. A varying number of experts is appointed in the Netherlands and France, while England and Sweden always rely on multiple experts (respectively two and three). In the Netherlands, mandatory forensic treatment under the Entrustment Act can only be imposed if there are at least two experts involved, among whom at least one psychiatrist. In all countries where more than one expert is involved, the assessment takes place under the supervision of a (forensic) psychiatrist. Germany, England, Sweden and Canada have certified training in forensic psychiatry (Nedopil, Gunn, & Thomson, 2012). In Germany however, this training is not mandatory to conduct forensic assessments. Other

Table 1. Main features of the assessment on criminal responsibility

	Forensic assessment mandatory	Moment assessment order	Criteria assessment order	Clinical observation	Format	Standard legal insanity
Belgium	No	Any time	No	No-	No	Cognitive and volitional
the Netherlands	No*	Any time	BooG	Yes	Guidelines	Volitional
France	Yes×	Any time	No	Yes	Guidelines	Cognitive and volitional
Germany	No*	Any time	No	Yes	Guidelines	Cognitive and volitional
England	Yes*	Any time	No	Yes	Legally defined + guidelines	Cognitive
Sweden	NA	Pre-sentence	Yes	Yes	Yes	No insanity defence
Canada	Yes	Any time	Yes	Yes	No	Cognitive

* Expert opinion is only mandatory in case compulsory treatment is to be imposed.
× Mandatory in case of severe offenses or sexual offenses.
- Although legal possibility.

countries also provide some sort of forensic training, although not formalized (Nedopil et al., 2012). Increasingly, it is an expectation of the court that the assessing psychiatrist must have the requisite training and/or expertise to conduct forensic evaluations in accordance with professional guidelines.

With respect to the financial remuneration, considerable differences between the countries can be found. Furthermore, the fees are often dependent on the type of assessment and/or the professional background of the expert. Estimations in case of criminal responsibility evaluations are presented in Table 2. In Sweden, a governmental agency is responsible for the assessment. Besides their regular salary, staff members do not receive any additional fee for the assessments.

RECOMMENDATIONS

To warrant qualitative and reliable assessments, a few recommendations can be made. First of all, psychiatrists generally base their opinions on records and interviews with the defendant and third parties. It is therefore advisable, at least for the more complicated cases, to include a psychologist as psychological tests provide an additional source of information and serve as a more valid and reliable foundation of the clinical findings (Lally, 2003). Static or dynamic scales for the assessment of the risk of recidivism may also be useful (Ducro & Pham, in press), as well as neuropsychological tests in case of mental retardation or psychosis.

Second, it is recommended that an independent organization be established to preserve the quality of the reports by recognizing and supervising experts. An example of this can be found in the Netherlands, with the Dutch Register Forensic Experts (*Nederlands Register Gerechtelijk Deskundingen, NRGD*) and the Dutch Institute for Forensic Psychiatry and Psychology (*Nederlands Instituut voor Forensische Psychiatrie en Psychologie, NIFP*), and in Sweden with the National Board of Forensic Medicine. Third, a standardized format for report writing enhances quality. As the proposed guidelines in the different countries look similar, it seems that there is some consensus regarding the ideal forensic psychiatric/psychological report. However, when (strict) criteria are applied with regard to the forensic report or expert, it is recommended that specialized training and appropriate financial compensations be provided.

Table 2. Main features forensic psychiatric experts assessing criminal responsibility

	Number of experts	Professional background experts	Additional criteria expert	Specialized training	Fee (€)°
Belgium	Not defined	Not defined×	No	No	374/case
the Netherlands	1–3*	Psychiatrist, (psychologist), (social worker)	Yes	Yes -	90–125/h
France	≥ 1*	Psychiatrist, (psychologist)	Yes	No	277,5-296/case
Germany	Not defined	Not defined×	No	Yes	75–100/h
England	2	Psychiatrist, any physician	Yes	Yes -	106–160/h
Sweden	3	Forensic psychiatrist, psychologist, forensic social worker	Yes	Yes	no additional fee
Canada	1	Psychiatrist	No	Yes	no additional fee

* Depending on type of assessment: the Netherlands: single, double, triple, clinical; France: one, two or three levels.
× In practice, this is done by a psychiatrist (whether or not assisted by a psychologist).
- Obligatory.
° Depending on type of assessment and professional background of the expert.

CONCLUSION

In sum, it appears that there are distinct differences between the abovementioned countries with respect to criminal responsibility assessments. Although Canada is considered a pioneer with regard to forensic mental health, England, the Netherlands and Sweden appear to have a well-established system in conducting these assessments. In Sweden the system is very strict, meaning that all reports are delivered by a governmental agency with their own staff. The court orders the report from the agency and not from the experts. At the other end of the spectrum, Belgium and France are less organized, with no criteria concerning the experts or format and clearly underpaid experts, albeit some legal changes are expected. The first prerequisite for the provision of quality forensic assessments is however, the presence of appropriate training (Every-Palmer et al., 2014).

REFERENCES

Bal, P., & Koenraadt, F. (2000). Criminal law and mentally ill offenders in comparative perspective. *Psychology, Crime & Law, 6*(4), 219-250. doi: 10.1080/10683160008409805

Belfrage, H., & Fransson, G. (2000). Swedish forensic psychiatry. A field in transition. *International Journal of Law and Psychiatry, 23*(5-6), 509-514.

Boetticher, A., Nedopil, N., Bosinski, H. A. G., & Saß, H. (2007). Mindestanforderungen für Schuldfähigkeitsgutachten. *Forensische Psychiatrie, Psychologie, Kriminologie, 1*(1), 3-9. doi: 10.1007/s11757-006-0002-8

Dressing, H., & Salize, H. J. (2006). Forensic psychiatric assessment in European Union member states. *Acta Psychiatrica Scandinavica, 114*(4), 282-289. doi: 10.1111/j.1600-0447.2006.00771.x

Ducro, C., & Pham, H. T. (in press). Recidivism risk assessment in Belgium: Clinical practice, research, and legislative aspects. In J. P. Singh, S. Bjorkly, & S. Fazel (Eds.), *International perspectives on violence risk assessment*. New York, NY: Oxford University Press.

Every-Palmer, S., Brink, J., Chern, T. P., Choi, W. K., Hern-Yee, J. G., Green, B., . . . Mellsop, G. (2014). Review of psychiatric services to mentally disordered offenders around the Pacific Rim. *Asia-Pacific Psychiatry, 6*(1), 1-17. doi: 10.1111/appy.12109

Fédération Française de Psychiatrie. (2007). *Expertise psychiatrique pénale: Audition publique 25 et 26 janvier 2007*. Retrieved from http://www.has-sante.fr/portail/upload/docs/application/pdf/epp-recommandations_de_la_commission-version_finale.pdf

Glancy, G. D., Ash, P., Bath, E. P. J., Buchanan, A., Fedoroff, P., Frierson, R. L., . . . Zonana, H. V. (2015). AAPL practice guideline for the forensic assessment. *Journal of the American Academy of Psychiatry and the Law, 42*(Suppl. 2), S3-S53.

Kordelaar, W. F. Van (2002). *Beslissingsondersteuning onderzoek Geestvermogens in het strafrecht voor volwassenen*. Deventer: Gouda Quint.

Lally, S. J. (2003). What tests are acceptable for use in forensic evaluations? A survey of experts. *Professional Psychology: Research and Practice, 34*(5), 491-498. doi: 10.1037/0735-7028.34.5.491

McLeod, R., Sweeting, A., Joyce, L., Evans, R., & Barkley, C. (2010). *Good practice guidance: Commissioning, administering and producing psychiatric reports for sentencing*. Retrieved from http://www.ohrn.nhs.uk/resource/policy/GoodPracticeGuidePsychReports.pdf

Nederlands Instituut voor Forensische Psychiatrie en Psychologie. (2014). *Schrijfwijzer psychiatrisch en psychologisch onderzoek pro justitia*. Retrieved from https://www.nifpnet.nl/NIFP/Beoordelen/Formatsstraf.aspx

Nedopil, N., Gunn, J., & Thomson, L. (2012). Teaching forensic psychiatry in Europe. *Criminal Behaviour and Mental Health, 22*(4), 238-246. doi: 10.1002/cbm.1845

Wettstein, R. M. (2005). Quality and quality improvement in forensic mental health evaluations. *Journal of the American Academy of Psychiatry and the Law, 33*(2), 158-175.

World Psychiatric Association. (2015). *Consensus guidelines for independent medical examinations*. Retrieved from http://wpanet.org/detail.php?section_id=11&content_id=472

Demographic, clinical, and risk characteristics of the study population

3

INTRODUCTION

In 2001, three forensic psychiatric units with ten beds each were established by the Minister of Justice in the Public Psychiatric Care Center in Rekem, the psychiatric center Sint-Jan-Baptist in Zelzate, and the university psychiatric center Sint-Kamillus in Bierbeek. In 2001–2002, the Minister of Health added three additional small units, with eight beds each, in the same hospitals (De Vuysere, 2005). In 2007, the units of the Department of Health and Justice were merged and are now referred to as medium security units (MSUs), and lie under the jurisdiction of the Ministry of Health. Currently, there are 193 medium security beds in Belgium. In Flanders the beds are equally divided among the three projects (45 beds each), while in Wallonia the beds can be found in the psychiatric hospital Jean Titeca in Brussels (22 beds), the psychiatric hospital Les Maronniers in Tournai (16 beds), and the neuropsychiatric center Saint-Martin in Dave (20 beds) (De Rycke & Pauwelyn, 2015). What started as a small project in Flanders has grown into a differentiated Flemish categorical network of forensic services and institutions, including outreach teams, forensic sheltered homes, and forensic psychiatric care centers (forCare; *forensisch psychiatrisch verzorgingstehuis*, forPVT) for the continued treatment and care after medium security treatment. Besides the medium security network, the government created several other initiatives by establishing, for instance, units for sex offenders and closed forensic protected houses for offenders with an intellectual disability (for an overview, please see De Rycke & Pauwelyn, 2015). Having implemented the first forensic psychiatric beds in Flanders, the medium security units can be considered pioneers in the forensic field.

The study population in this dissertation consists of patients admitted during the first 10 years after the establishment of the MSUs, which provide treatment for medium risk and medium security internees (i.e., average safety and hazard risk internees). Only the forensic psychiatric hospital in Zelzate provides treatment for female internees, and has been doing so since 2006. The treatment at MSUs is meant for Dutch-speaking internees who have some degree of motivation and learning ability and preferably a psychotic and/or personality disorder. Exclusion criteria for admission are primary sexual or addiction problems and/or high levels of psychopathy (De Smedt, Mariën, & Vermeiren, 2008). A description of the treatment programs within these units is beyond the scope of this dissertation and can be found in Boers,

Vandevelde, Soyez, De Smet and To (2011). In general, forensic psychiatric treatment differs substantially from treatment in general psychiatry, in that it addresses both the mental disorder and the criminal or violent propensity. Besides the duality of treatment and control, there is more emphasis on the importance of assertive and monitored aftercare and the enhancement of motivation, which is generally poor to non-existent at first. All medium security units use cognitive behavioral therapeutic approaches within a relapse prevention framework. The overall treatment aim is primarily to reduce risk levels in order to transfer the patient to a lower-security level. Following the Risk-Need-Responsivity (RNR) principles (Andrews, Bonta, & Hoge, 1990), treatment focuses on criminogenic needs and the intensity of the treatment is related to the risk level. Also, responsivity issues, such as the degree of connection in the treatment, is addressed (Schuringa, Spreen, & Bogaerts, 2014).

In this chapter, some of the basic concepts will be defined, such as risk and security, because of confusion about their interpretation. Additionally, it is interesting to note that the name of the population was changed from *medium risk* internees to *medium security* internees.

Defining medium security

At the start of the MSUs, an interuniversity research project was ordered to evaluate the extent to which a medium risk/security group differs from a low and high risk group of internees (Casselman et al., 2002). A theoretical model to describe the three security levels, which is related to the RNR model, was proposed by Vertommen and Maesschalck (2007) as part of this project and was largely based on conceptual and operational definitions used in the UK (Table 1).

Table 1. Conceptual model of risk and security level

	Patient perspective	Unit perspective
Level 1	Risk level	Security level
Level 2	Treatability	Care/cure/(observation)
Level 3	Psychopathology	Treatment programs

The first level relates to the risk and security level. *Risk* is the degree of violent propensity as well as the risk of (violent) reoffending, which can be low, medium, or high, and is usually measured through the use of structured risk assessment tools. Risk refers to the first principle (*risk principle*) of the RNR model, stating that the most intensive services should be delivered to the highest risk offenders (Andrews et al., 1990). Risk is a dynamic concept that changes over time, situation, and context. Therefore, translating the risk principle into an operational definition is difficult. On the contrary, *security* refers to the infrastructural level of security of an institution or unit (low, medium, or high), which is static in nature and related to the risk level a patient poses. A patient can have low, medium, or high security *needs*, while a unit provides a security *level*. The security level can be described according to its environmental (e.g., escape-proof building in the case of medium security), procedural (e.g., observed visits and limited access to cash in the case of medium security), and relational security characteristics (Kennedy, 2002). Relational security is closest to quality of care and includes – in the case of medium security – high patient-to-staff ratios, intensive treatment programs, and specific therapeutic skills to deal with dangerous, severely disturbed mentally ill patients. As such, it is related to the second principle of the RNR-model (*need principle*), stating that criminogenic needs should be the treatment targets (Andrews et al., 1990).

The second level refers to the treatability of patients and includes psychopathology, motivation, and possibility of cooperating with staff. Depending on treatability, care (i.e., support and guidance) or cure (i.e., treatment and therapy) is provided, or in case of doubt a period of observation is used. While care is completely independent from security level and can thus be provided at each security level, the delivery of cure or treatment can be hampered by medium and high security levels. At the third level, six separate treatment programs for specific psychopathologies were identified, including separate programs for psychotic, personality disordered, psychopathic, and intellectually disabled internees, as well as programs for sexual delinquents and women. The second and third level are related to the third principle of the RNR-model (*responsivity principle*), stating that treatment should be provided in a style and mode that is responsive to the offender's learning style and ability (Andrews et al., 1990).

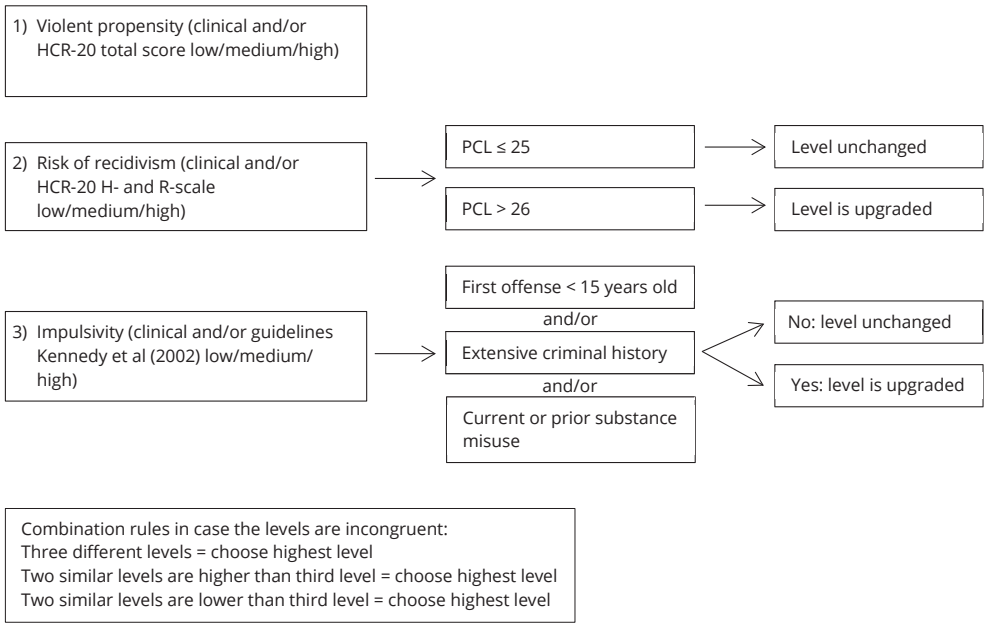


Figure 1. Flowchart to determine security level.

Starting from this general theoretical frame, Vertommen and Maesschalck (2007) further tried to operationalize the term medium security and defined criteria, which are shown in Figure 1. First, the risk and security level should be determined through 1) the propensity for violence (clinical assessment), 2) the risk of recidivism (assessed with the Historical, Clinical, Risk management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997), possibly combined with the Psychopathy Checklist-Revised (PCL-R; Hare, 2003), and 3) the presence of impulsivity (prior and/or current). Second, whether care, cure, or observation should be delivered should be assessed through 1) motivation, 2) alliance, 3) level of distress, and 4) psychopathology. Third, the need for a specific program should depend on the psychiatric diagnosis (Vertommen & Maesschalck, 2007). In addition, complementary to risk assessment instruments, different instruments are currently available in the literature to measure security level, such as the *Security Needs Assessment Profile* (SNAP; Collins & Davies, 2005) and the *DUNDRUM-1 Triage Security Scale* (Flynn, O'Neill, McInerney, & Kennedy, 2011). Notwithstanding the abovementioned theoretical

distinctions between security levels, according to Casselman et al. (2002) the indication for medium security admission was ultimately done in a "pragmatic, insufficient and sometimes even unacceptable way" (p. 19), and with a "severe lack of uniformity" (p. 25). In practice, clinicians acted as gatekeepers to MSU admissions and decisions were made on the basis of clinical judgment, with a panel of clinicians examining the ideal placement.

Previous research on the interned population in Flanders

To date, no studies other than the research included in this dissertation have been published with regard to the medium security population. One unpublished study, which yielded fewer results than originally expected according to Casselman (2011), will not be discussed, and took into account a small number of studied Flemish internees (*n* = 16). Previous research investigating the interned population in general noted a lack of systematic registration and scientific research (e.g., Casselman, Devuysere, & Vervaeke, 2003a). Most studies were performed on incarcerated internees (Cosyns, Dillen, de Ruyter, & De Doncker, 1994; Cosyns, 2005; De Vuysere, Casselman, & Vervaeke, 2004; Deckers et al., 2014; Vanden Hende, Caris, & De Block-Bury, 2005). Studies examining characteristics in the entire population were hampered because research was (partly) collected from unpublished reports (e.g., Goethals, 1985), covered small samples (e.g., Vandenbroucke, 1995), or assessed only one or a few variables such as intelligence (e.g., Maes, Goethals, & Verlinden, 2009) or offenses (e.g., Casselman, Devuysere, & Vervaeke, G., 2003b). No study was found that specifically examined internees residing in the community with ambulatory care. Some authors acknowledged that their findings only yielded rough indications, for instance of psychopathology (Deckers et al., 2014). Therefore, the summarized findings provided below are sometimes based on fragmented research and methodological shortcomings.

Socio-demographic characteristics indicated that internees were mostly unmarried (66.0%-85.4%), lesser educated males (3.0% to 7.3%), of Belgian nationality (84.7%), between 26 and 45 years old, and who were professionally inactive (47.9% to 77.7%) at the moment time of their internment (Casselman et al., 2003b; Cosyns, 2005; De Vuysere et al., 2004; Deckers et al., 2014; Goethals, 1985). De Vuysere et al. (2004) found that a quarter of incarcerated internees participated in special education and only 15.5% of the internees finished secondary school. In

the study by Goethals (1985), more than half only finished primary school. Previous admissions to a general psychiatric hospital were found in more than half (54.4%) of the internees (De Vuysere et al., 2004).

Older research found that most internment measures were for property offenses (Goethals, 1985), while more recent studies found that most common index offenses were violent offenses (Casselman et al., 2003b; Cosyns et al., 1994). The most serious index offenses, i.e., (attempted) manslaughter and homicide, were found in 3.0% to 10.5% of the population (Casselman et al., 2003b; Cosyns et al., 1994; Goethals, 1985). Judicial antecedents were not collected systematically in CPS Ghent and CPS Leuven, which resulted in missing information in one out of five files; 38.5% of the population was previously convicted and 16.6% had been previously interned (Casselman et al., 2003b).[1] When only valid percentages were taken into account, one out of four internees was a first-time offender (Casselman et al., 2003b), while older research found higher percentages of first offenders, ranging from 31% to 51% (Goethals, 1985).

Regarding primary psychiatric diagnoses, an older study found that intellectual deficiency was the most frequent diagnosis (Goethals, 1985), while more recent studies found that personality disorders were the most prevalent diagnoses, followed by psychotic disorder or intellectual deficiency (Cosyns et al., 1994; De Vuysere et al., 2004; Deckers et al., 2014; Vandenbroucke, 1995). Differences were noted when comparing incarcerated and conditionally released internees residing in psychiatric hospitals. In the latter, the most frequent diagnosis was psychotic disorder (42% as compared to 17% in incarcerated internees), while the most frequent diagnosis for incarcerated internees was personality disorder (42% as compared to 11% in hospitalized internees) (Cosyns, 2005). A primary diagnosis of paraphilic disorder was also more prevalent in incarcerated internees compared to hospitalized internees (17% vs. 10%; Cosyns, 2005). Dual or triple diagnoses were found in three quarters of the internees (Cosyns et al., 2007; Deckers et al., 2014). Cumulatively, 55.2% to 81.0% of personality disorders were found (Cosyns et al., 2007; Deckers et al., 2014; Vandenbroucke, 1995). Substance use disorders were found in 35.7% to 44.7% of the internees (Cosyns et al., 2007; Deckers et al., 2014; De Vuysere et al., 2004). Cumulatively, paraphilic disorders were found in 28.8% of the incarcerated population (Cosyns et al., 2007).

Intelligence scores were investigated in two studies, focusing on the incarcerated internees (Vanden Hende et al., 2005), and on the total interned population (Maes et al., 2009). An IQ score was lacking in 8.2% and 8.6% of the files (Maes et al., 2009; Vanden Hende et al., 2005). Intellectual disability was found in 21.8%[2] and 18.5%[3] of the population, and borderline intellectual functioning was found in 19.2% and 14.7% of the population (Maes et al., 2009; Vanden Hende et al., 2005). Surprisingly, the exact intelligence tests used were not specified in these studies. This might however be important since substantial dissimilarities between tests exist, according to Habets, Jeandarme, Uzieblo, Oei and Bogaerts (2014). They showed that 33% to 66% of cases had more than a 10-point difference in reported IQ score, depending on the test used, for a medium security population (Habets et al., 2014).

International profile of patients admitted in (medium security) forensic units

Demographically, these patients were mostly single, poorly educated, unemployed men between 30 and 40 years old (Blattner & Dolan, 2009; Coid, Kahtan, Gault, Cook, & Jarman, 2001; Dolan & Khawaja, 2004; Gow, Choo, Darjee, Gould, & Steele, 2010; Lelliott, Audini, & Duffett, 2001; Melzer et al., 2004). With the exception of a study of an inner-London population (Lelliott et al., 2001), the majority of the patients were Caucasian (Blattner & Dolan, 2009; Coid et al., 2001; Dolan & Khawaja, 2004). The majority (> 75%) had previous admissions to a general psychiatric hospital (Blattner & Dolan, 2009; Gow et al., 2010; Lelliott et al., 2001; Linhorst & Scott, 2004; Melzer et al., 2004). A quarter (22%; Coid et al., 2001) to half (47%; Dolan & Khawaja, 2004) of the population was admitted on more than one occasion to a MSU. After medium security treatment, approximately half of the population was discharged to the community (Blattner & Dolan, 2009; Gow et al., 2010). In the study of Dolan and Khawaja (2004), the majority was discharged to the community with forensic community care.

Judicially, most patients were not first-time offenders but recidivists (Blattner & Dolan, 2009; Dolan & Khawaja, 2004; Freestone et al., 2012; Gow et al., 2010; Lelliott et al., 2001; Melzer et al., 2004). Linhorst and Scott (2004) found that 36.7% had prior convictions for serious crimes (i.e., felony convictions). Index offenses were mostly violent offenses (e.g., manslaughter or battery), followed by property offenses (e.g., theft and arson) and sexual offenses (Blattner &

Dolan, 2009; Coid et al., 2001; Dolan & Khawaja, 2004; Freestone et al., 2012; Gow et al., 2010; Lelliott et al., 2001).

Most primary psychiatric diagnoses were psychotic disorders (mostly around 60% to 70%) and personality disorders (around 10% to 30%; Blattner & Dolan, 2009; Coid et al., 2001; Dolan & Khawaja, 2004; Gow et al., 2010; Ibishi, Musliu, Hundozi, & Citaku, 2015; Lelliott et al., 2001; Melzer et al., 2004). Most studies found comorbid substance misuse in up to half of the population (Blattner & Dolan, 2009; Carr et al., 2006; Dolan & Khawaja, 2004; Gow et al., 2010; Gradillas, Williams, Walsh, & Fahy, 2007; Hildebrand & de Ruiter, 2004). Overall, high rates of comorbidity were found (Blattner & Dolan, 2009; Gow et al., 2010). Among personality disorders, cluster B personality disorders were the most prevalent disorders found (Hildebrand & de Ruiter, 2004; Pham & Saloppé, 2010; Spreen, Brand, Ter Horst, & Bogaerts, 2014; Timmerman & Emmelkamp, 2001). Furthermore, frequent comorbidity was found between borderline personality disorder and antisocial (APD) personality disorder, with the highest prevalence for APD in a Dangerous and Severe Personality Disorder (DSPD) medium security unit in the UK (Freestone et al., 2012). In addition, APD was the most frequent diagnosis in other medium security and high security units (Dolan & Khawaja, 2004; Hildebrand & de Ruiter, 2004; Pham & Saloppé, 2010; Timmerman & Emmelkamp, 2001). Contrary to other studies, Hildebrand and de Ruiter (2004) and Timmerman and Emmelkamp (2001) found a higher number of personality disordered patients (89%) than patients with a psychotic disorder (19%).

In non-Western countries, similar demographic, clinical, and judicial characteristics were found in forensic psychiatric populations. The majority of the patients were young, unemployed, poorly educated single men with a psychotic disorder and substance misuse who were being treated for a violent index offense (Barrett et al., 2007; Pal, 1997; Yusuf & Nuhu, 2009). Epilepsy and other organic disorders were noted in about 10% of all diagnoses (Pal, 1997; Strydom, Pienaar, Dreyer, van der Merwe, & Jansen van Rensburg, 2011). Among the violent index offenses, a high number of rapes were noted (Barrett et al., 2007; Strydom et al., 2011). Menezes, Oyebode, and Haque (2009) and Yusuf and Nuhu (2009) noted a very high number (68% to 71%) of homicidal index offenses, which was explained by the fact that in Africa, many patients suffering from a major mental illness remained untreated in the community and came to the attention of the psychiatric services only after they committed an offense (Menezes et al., 2009).

Mean HCR-20 scores in medium security populations in the UK ranged from 18.3 to 20.5 and were higher in personality disordered patients (*M* = 26.1; Dolan & Khawaja, 2004; Freestone et al., 2012; Gray et al., 2004; Gray, Taylor, & Snowden, 2008). Regarding PCL-R total scores, the mean score in a validation sample of forensic psychiatric patients was 21.5 (Hare, 2003). Personality disordered medium security patients in the UK had significantly higher PCL-R total scores (*M* = 23.4) than a representative prison population (*M* = 14.8; Freestone et al., 2012). In a Dutch study, male TBS-patients had a mean PCL-R total score of 21.4 (Hildebrand & de Ruiter, 2004). In Belgium, mean PCL-R total scores in internees were 17.6 and 19.6, and these scores were comparable to mean scores in convicts (*M* = 17.5; Pham, Saloppé, Bongaerts, & Hoebanx, 2007; Pham & Saloppé, 2010). The mean Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice, & Cormier, 2006) total score in the validation sample of offenders undergoing a forensic psychiatric evaluation was 0.9. In other forensic psychiatric studies, the mean VRAG total scores ranged from 3.0 to 9.7 (Gray, Fitzgerald, Taylor, MacCulloch, & Snowden, 2007; Ho, Thomson, & Darjee, 2009; Snowden, Gray, & Taylor, 2010; Snowden, Gray, Taylor, & Fitzgerald, 2009; Snowden, Gray, Taylor, & MacCulloch, 2007).

The variable prevalence rates can be explained by several factors. They may reflect organizational aspects of the forensic care system (Salize & Dressing, 2007). For example, TBS-order under the Dutch Entrustment Act is only intended for mentally ill offenders who have committed a serious violent offense, and can also be imposed on offenders with only diminished responsibility (Hildebrand & de Ruiter, 2004). Furthermore, it should be acknowledged that levels of security are not necessarily identified in all countries. Countries such as the UK, who acknowledge medium security levels, provide medium security treatment for several patient groups: those who no longer require treatment in a high security hospital, those admitted from prison or the court, and difficult-to-manage patients cared for by general psychiatric services (Melzer et al., 2004). As noted by Rice and Harris (1997), the term *mentally disordered offender* encompasses a heterogeneous and poorly defined group, including insanity acquittees, persons found guilty but mentally ill, persons found unfit to stand trial, mentally disordered sex offenders, sexual predators, and prisoners transferred to mental health facilities, all of whom have a wide range of treatment needs. In addition, the admissions policy of local institutions with different inclusion and exclusion criteria affects the profile of the study populations (e.g., DSPD

unit Freestone et al., 2012). Furthermore, Coid and Kathan (2000) have argued that other factors, such as notoriety of the index offense and absence of alternative treatment options, can play a role. Study methodology can influence population profiles, for instance, when severely disturbed or psychotic patients are excluded from the study (Pham & Saloppé, 2010; Timmerman & Emmelkamp, 2001). Finally, other methodological differences can influence study results and prevalence rates, depending on whether patient notes, self-report, or semi-structured interviews are used (Hildebrand & de Ruiter, 2004). In sum, generalizing from local studies can be difficult and there are very few national studies with which to compare results.

METHOD

Procedure

Information on medium security admissions in the period 2001–2010 was provided to the researchers by the three MSUs, as well as basic information on socio-demographic, clinical, and judicial variables. In order to ensure the validity of the data collected, all variables were uniformly re-coded by researchers after cross-referencing with all of the available information that could be found in hospital records and individual files of the CPS. Individual CPS files contain psychiatric reports on accountability, information on previous and current hospitalizations, as well as detention periods.

The following socio-demographic categorical variables were collected from the hospital and CPS files: sex of the offender, age (clustered as follows: under 24, 25–34, 35–44, 45–54, 55–64, and 65 or older), nationality, country of birth, type of education and highest level of education, and previous admissions to a general psychiatric unit prior to first medium security admission. Several variables were coded at the moment of the index offense: marital status, partner, children, and social competence. The System for the Classification of Extra Familial Pedosexuals (*Systeem voor de Classificatie van Extrafamiliale Pedoseksuelen (SCEP)*; De Doncker, Koeck, Huys, & Winter, 2006) was used to score five social competences at the time of the index offense according to the Massachusetts Treatment Centre Child Molester Typology, third version (MTC:CM3; Knight, Carter, & Prentky, 1989). The following SCEP-criteria were coded: 1) work (having worked at least three years in a stable job or a sheltered work environment, excluding

work in prison; Job changes counted if they referred to a promotion or were characteristic for the job), 2) relationship (was in an intimate relationship with an adult for at least one year, married, or living together), 3) children (took care of and was responsible for biological children or stepchildren for at least three years), 4) engagement (was an active member of a club or organization for at least one year, where engagement was defined as actively participating, taking responsibility in the board, etc.), and 5) friendship (was in a reciprocal friendship with an adult other than an intimate partner for at least one year, as characterized by mutual activities and not opportunistic in nature). A patient was considered highly socially competent if at least two SCEP-criteria were present. The following socio-demographic continuous variables were collected from hospital files: age at first admission and age at first discharge from MSU.

The following judicial categorical variables were collected from CPS files and extracts of the Central Criminal Record of the Ministry of Justice: nature of the index offense (offense for which the internment measure was imposed and which was the basis for the transfer to a MSU), juvenile court sentences, as well as presence and nature of prior sentences (further divided into prior convictions and internments). Recent (extracted by the author since February 16, 2012) and older Central Criminal Records (found in CPS files) were compared, and in case of discrepancies all of the sentences were counted. Discrepancies were found in 113 files (21.3%). In cases of multiple crimes, the offense was qualified according to the level of moral severity, based largely on the Crime Severity Score (DE-12; Brand, 2005). The following continuous variables were collected: age at index offense, number of prior sentences, convictions and internments for different categories of offenses, age at first sentence, conviction and internment, and number of different offense categories. Five offense categories were defined, as follows: non-violent sexual offenses, property or acquaintance offenses, drug-related offenses, other offenses, and violent offenses (including sexual violence and arson where persons were endangered). Violence referred to the intentional use of physical force or power – threatened, attempted, or actual – against another person. Verbal interpersonal violence referred to threatened violence (i.e., verbal violence); physical violence referred to attempted or actual interpersonal violence (i.e., physical violence). The number and duration of prior detention periods in prison were calculated on the basis of the detention records from the administrative prison registration system *Detentie Informatie Systeem (SIDIS),* extracted on November 16, 2011.

The following clinical categorical variables were collected from hospital and CPS files. All (comorbid) diagnoses were based on the Diagnostic and Statistical Manual of mental disorders-IV-Text Revision (DSM-IV-TR; American Psychiatric Association, 2000). Diagnoses were extracted from the Minimum Psychiatric Data (MPD) registration system (*Minimale Psychiatrische Gegevens*, *MPG*), which is a mandatory registration system used by psychiatrists in Belgian psychiatric hospitals. The MPD data were cross-referenced with relevant information found in the hospital and CPS files. Discrepancies were discussed with the treating psychiatrists and corrected as needed. This was the case for 46.7% of the diagnoses. Intelligence scores were based on scores obtained on the Wechsler Adult Intelligence Scale-III (Wechsler 2005), found in the files. In cases where there were multiple scores on file, the most recent WAIS-III score was used. The following psychopathy scores and risk assessment scores were collected: scores on the PCL-R, the VRAG, and the HCR-20. With the exception of the VRAG, which was scored on a convenience sample of 230 patients by the researchers, other risk assessment scores were field scores, and were coded prospectively as part of clinical practice. In case of multiple scores, the PCL-R and HCR-20 were based on the first assessment. If scored in a hospital setting, the scoring was conducted by criminologists or psychologists. If scored in a prison setting, scoring was performed by prison psychologists or psychologists/psychiatrists conducting a pre-trial insanity evaluation.

Ethical considerations

Since this was a retrospective file study, informed consent from the patients was not obligatory. However, one local ethical committee requested that information letters be sent to all the internees, explaining the purpose of the study and the variables that were collected. As a result, 11 internees refused to give passive consent and were therefore excluded from the analyses, resulting in a study population of 531 internees.

The study was approved by the Ethics committee of Antwerp University Hospital on January 24, 2011.

Data analysis

Simple descriptive analyses were conducted using the software package SPSS version 22 (IBM Corp, Released 2013). Valid percentages are given. Since the data sources were characterized by different percentages of missing data, some analyses were based on smaller samples, as is noted throughout the chapter and in the tables. The difference between performal and verbal IQ was analyzed with a Wilcoxon Signed Ranks Test.

RESULTS

Socio-demographics

Socio-demographic characteristics for the entire population (*N* = 531) are presented in Table 2. The study sample was predominantly male (94.9%, *n* = 504) and had Belgian nationality (90.1%, *n* = 475). Most internees were born in Belgium (89.4%, *n* = 472); the other internees (10.6%, *n* = 56) were born in Africa (50%), Europe (39.3%), Asia (7.1%) and America (3.6%). The mean age at first admission in a MSU was 36.5 years (*SD* = 10.82, range = 18.8–73.4 year). A minority (12.8%, *n* = 68) was younger than 25; few internees (1.7%, *n* = 9) were 65 years or older. The most common age category was between 25 to 34 years old (38.4%, *n* = 204). Marital status, children and intimate partner relationships were assessed at the time of the index offense. Most (84.6%, *n* = 435) were living alone (single or separated). A quarter (27.3%, *n* = 131) was in a romantic relationship and almost a third (29.1%, *n* = 152) had biological children. Education was assessed in two ways: type of education and, in case of normal education, level of achieved education. Most participants followed normal education (75.2%, *n* = 373), while a quarter (23.6%, *n* = 117) followed special education. In addition, six internees followed no education. In the group with normal education, only a quarter actually finished high school with a degree[4] (24.5%, *n* = 86), while most internees had some degree of high school education (68.4%, *n* = 240). In the group with high school diploma, 16 internees had a continuing (university) training. Most participants (81.6%, *n* = 425) had been previously admitted to a general psychiatric unit, either as a voluntary patient or civilly committed. Social competence variables were analyzed up until the first admission. A stable work history was found in a third (30.6%, *n* = 148) of the population. A stable relationship history was

Table 2. Socio-demographic characteristics

			Population assessed			
	N	Missing (%)	*n*	%	*M*	*SD*
Sex	531	0				
Male			504	94.9		
Female			27	5.1		
Age category	531	0				
< 25			68	12.8		
25–34			204	38.4		
35–44			144	27.1		
45–54			85	16.0		
55–64			21	4.0		
≥ 65			9	1.7		
Belgian nationality	527	0.8	475	90.1		
Born in Belgium	528	0.6	472	89.4		
Previous admissions	521	1.9	425	81.6		
Marital status	514	3.2				
Married/living common-law			79	15.4		
Single/separated			435	84.6		
Partner at index offense	480	9.6	131	27.3		
Children at index offense	523	1.5				
0			371	70.9		
≥ 1			152	29.1		
1			74/152	48.7		
2-3			67/152	44.1		
≥ 4			11/152	7.2		
SCEP criteria						
Low social competence	135	74.6	96	71.1		
High social competence	135	74.6	39	28.9		
Single SCEP variables						
Stable work history ≥ 3 y.	484	8.9	148	30.6		
Stable relationship history ≥ 1 y.	474	10.7	204	43.0		
Nurturing children ≥ 3 y.	469	11.7	55	11.7		
Active member organization ≥ 1 y.	285	46.3	36	12.6		
Active friendship ≥ 1 y.	189	64.4	39	20.6		
Education	496	6.6				
Regular education			373	75.2		
Special education			117	23.6		
No education			6	1.2		
Regular education	351	5.9				
Degree high school			86	24.5		
Degree university/higher education			16	4.6		
Age at admission	531	0			36.5	10.82
Age at first discharge	502				37.8	10.87

found in 43.0% (*n* = 204) of the population. A minority (11.7%, *n* = 55) took responsibility for a child. For the two remaining variables related to social competence, there were a lot of missing data. In 342 internees (64.4%) there was no information concerning stable friendships and in 246 internees (46.3%) there was no information on organization membership. A minority (20.6%, *n* = 39) had a stable friendship and a minority (12.6%, *n* = 36) was an active member of a leisure organization of peers at least during one year prior to the index offense. In only a quarter of the cases (25.4%, *n* = 135) all five SCEP criteria were available. In this subpopulation, a quarter (28.9%, *n* = 39) could be considered high socially competent.

Taken together, the prototypical medium security patient was a Belgian male between 25 and 34 years old, without a high school diploma and with an unstable work and relationship history, who was living alone and having no partner or children at the time of the index offense.

Criminal justice involvement

Index offense

The index offenses are presented in Table 3. The index offense was characterized as the primary offense leading to the internment measure, which was the basis for the referral to a medium security unit. The average age at the index offense was 32.1 years (*SD* = 10.01, range = 18.5–70.3). On average, there were 2.5 (*SD* = 2.05, range = 1–14) offenses at the index internment measure. If multiple offenses were present, the index offense was coded by the most serious offense. Participants were found NGRI for a variety of offenses: violent offenses (77.2%, *n* = 410), property offenses (18.6%, *n* = 99), drug-related offenses (2.3%, *n* = 12), sexual non-violent offenses (0.8%, *n* = 4) and other offenses (1.1%, *n* = 6). Almost a fifth (18.1%, *n* = 96) committed an attempted (9.0%, *n* = 48) or actual manslaughter or homicide (9.0%, *n* = 48). A minority (7.5%, *n* = 40) committed a sexual hands-on offense. Adding sexual hands-on and hands-off offenses, 44 internees (8.3%)committed a sexual offense. Forty-three internees (8.1%) were found NGRI after arson, either in order to destroy property or to intentionally or recklessly endanger life.

Table 3. Index offense

		Total population			
	N	*n*	%	*M*	*SD*
Index offense	531				
Violence		410	77.2		
Violence other		145	27.3	145/410	35.4%
Property with violence		62	11.7	62/410	15.1%
Manslaughter/homicide		48	9.0		
Attempted manslaughter/homicide		48	9.0		
Verbal violence		37	7.0		
Arson intentionally or recklessly endangering life		30	5.6		
Sexual hands-on adult		21	4.0		
Sexual hands-on minor		19	3.6		
Property		99	18.6		
Arson to destroy property		13	2.4		
Property offense e.g., theft		86	16.2		
Drugs		12	2.3		
Sexual hands-off		4	0.8		
Other		6	1.1		
Traffic		2	0.4		
Not otherwise specified		4	0.8		
Age at index offense	531			32.1	10.01

Judicial history (Table 4)

Most participants had prior contact with the criminal justice system as a juvenile or adult (84.4%, *n* = 448); only a minority (15.6%, *n* = 83) was a first offender. Four out of ten internees (40.6%, *n* = 204) were convicted in juvenile court. In 28 cases the reason was not clear; in the other cases (*n* = 176) the majority of patients had been subjected to juvenile court (86.9%, *n* = 153) after having committed an offense *(als Misdrijf Omschreven Feit, MOF),* either alone (*n* = 123) or in combination (*n* = 30) with a problematic parenting situation *(Problematische Opvoeding Situatie, POS)*. The mean age at first conviction or internment – index offense included – was 24.8 years (*SD* = 9.13, range = 9.8–69.1). The mean age at first conviction (*n* = 425) was 23.3 years (*SD* = 7.81, range = 9.8–61.8) and at first internment (*n* = 531) 30.7 years (*SD* = 9.71, range = 17.6–69.3). One internee was

Table 4. Judicial history

			Population assessed			
	N	Missing (%)	*n*	%	*M*	*SD*
Juvenile court	503	5.3	204	40.6		
MOF and POS known	176	13.7				
MOF			123	69.9		
MOF and POS			30	17.0		
POS			23	13.1		
Prior sentences total	531	0	448	84.4		
0			83	15.6		
1			42	7.9		
2-3			64	12.1		
≥ 4			342	64.4		
Prior sentence violent offense	531	0	341	64.2		
Number prior sentences	448				6.2	5.73
Number violent sentences	341				2.5	180
Prior convictions	531	0	425	80.0		
0			106	20.0		
1			44	8.3		
2-3			81	15.3		
≥ 4			300	56.5		
Prior convictions violent offense	531	0	299	56.3		
Prior convictions (attempted) manslaughter/homicide	531	0	11	2.1		
Prior convictions sexual offense	531	0	40	7.5		
Number prior convictions	425				5.8	5.38
Number prior convictions violence	299				2.2	1.60
Age first conviction	425				23.3	7.81
Prior internments	531	0	188	35.4		
0			343	64.6		
1			43	8.1		
2-3			61	11.5		
≥ 4			84	15.8		
Prior internments violent offense	531	0	135	25.4		
Prior internments (attempted) manslaughter/homicide	531	0	15	2.8		
Prior internments sexual offense	531	0	20	3.8		
Number internments	188				1.7	1.22
Number internments violence	135				1.4	0.70
Age first internment	531	0			30.7	9.71
Antecedents index included	531	0			6.3	5.73
Antecedents index included violence	531	0			2.4	1.91
Age first sentence antecedents index incl.	531	0			24.8	9.13
Detention periods	530	0.2	526	99.2		
Number of prior detentions	530	0.2			4.2	4.36
Duration of prior detentions (years)	530	0.2			4.6	4.99

Note. MOF = *Misdrijf Omschreven Feit* (juvenile crime); POS = *Problematische Opvoedingssituatie* (problematic parenting situation); sentence = conviction or internment measure.

younger than 18 at the first internment verdict. In the 448 internees (84.4%) who had been sentenced prior to the index offense, on average 6.2 (*SD* = 5.73, range = 1–39) prior sentences were found and in the 341 internees (64.2%) who had been sentenced for a violent offense prior to the index offense, on average 2.5 (*SD* = 1.80, range = 1–10) prior violent sentences were found. In Table 4 prior sentences are further divided into prior convictions (80%, *n* = 425) and prior internment measures (35.4%, *n* = 188). In addition, when adding prior sentences to the index internment measure, the mean amount of sentences for a general offense was 6.2 (*SD* = 5.73, range = 1–40), and for a violent offense, it was 2.4 (*SD* = 1.91, range = 0–11). On average internees committed 2.8 different categories of offenses (*SD* = 1.07, range = 1–5).

The majority of the population (99.2%, *n* = 526) was incarcerated in prison prior to the medium security admission. On average internees were incarcerated 4.2 times (*SD* = 4.36, range = 0–37), in total for a duration of 4.6 years (*SD* = 4.99, range = 0–37).

Taken together, the prototypical medium security patient was interned for a violent index offense, was on average 24.8 years old when first sentenced, had 6.3 sentences, mostly convictions and was incarcerated 4.2 times in prison for a total duration of detention of 4.6 years prior to first admission in medium security.

Clinical diagnoses

DSM-IV-TR cumulative diagnoses

Each internee suffered from an Axis I and/or an Axis II DSM-IV-TR diagnosis. Most (93.4%, *n* = 496) had a diagnosis on Axis I. Also, the majority (77.8%, *n* = 413) had a diagnosis on Axis II. Levels of comorbidity were high, with 378 internees (71.2%) combining Axis I and Axis II pathology. Also, comorbidity of Axis I combined with Axis II personality disorders was high (64.2%, *n* = 341).

Psychiatric diagnoses are first presented cumulatively in Tables 5 to 9, i.e., depending on whether the diagnoses was present either as a primary or as a comorbid diagnosis. The most common cumulative DSM-IV-TR diagnoses were personality disorders (70.6%, *n* = 375), substance use disorders (56.7%, *n* = 301), psychotic disorders (43.9%, *n* = 233) and intellectual disabilities (23.0%, *n* = 122).

In Table 5, more information is provided regarding the most commonly found diagnosis – personality disorder – which was the sole diagnosis in only 23 cases (6.1%). Most personality

disorders were comorbid disorders (93.9%, *n* = 352). Cluster B personality disorders were most frequently found. More specifically, most frequent diagnoses were antisocial personality disorder (*n* = 132), borderline personality disorder (*n* = 96) and personality disorder not otherwise specified (*n* = 83). Twenty internees had more than one personality disorder.

Table 5. Personality disorders

		Total population		Population with PD	
	N	*n**	%*	*n*	%
Any personality disorder	531	375	70.6		
Cluster A		36	6.8		
Paranoid		15	4.0		
Schizotypal		11	2.9		
Schizoid		10	2.7		
Cluster B		242	45.6		
Antisocial		132	35.2		
Borderline		96	25.6		
Narcissistic		24	6.4		
Histrionic		6	1.6		
Cluster C		17	3.2		
Avoidant		8	2.1		
Dependent		8	2.1		
Obsessive-compulsive		2	0.5		
Personality disorder NOS		83	22.1		
More than one personality disorder		20	5.3		
Comorbidity				375	100
Sole diagnosis		23		23/375	6.1
Combi Axis I and Axis II intellectual disability		79		79/375	21.1
Combi Axis I or Axis II intellectual disability		273		273/375	72.8
Combi Axis I		268			
Combi Axis II intellectual disability		5			

Note. Personality disorder NOS = personality disorder not otherwise specified.
* More than one cluster and/or personality disorder is possible.

Table 6 specifies the second most common diagnosis – substance misuse – which was found in more than half of the population (56.7%, *n* = 301). Of the 301 patients with substance misuse, there were 81 patients with alcohol problems (15.3% of the entire population), 59

patients with problems with another substance (11.1% of the entire population) and 161 patients with two or more substance problems (30.3% of the entire population). Taken together, alcohol misuse was found in 19.6% of the population (*n* = 104), either in isolation (*n* = 81) or in combination with another substance (*n* = 23). Substance use disorders were in most cases comorbid disorders (97.3%, *n* = 293), whereas it was the sole diagnosis in eight patients.

Table 6. Substance misuse

		Total population		Population with SUD	
	N	*n*	%	*n*	%
Any substance misuse	531	301	56.7	301	100
Alcohol		81	15.3	81/301	26.9
Other substance		59	11.1	59/301	19.6
≥ 2 substance misuse		161	30.3	161/301	53.5
Comorbidity				301	100
Sole diagnosis		8		8/301	2.7
Combi Axis I and II		137		137/301	45.5
Combi Axis I or II		156		156/301	51.8
Combi Axis I		46			
Combi Axis II		110			

Note. SUD = substance use disorder.

In Table 7 an analysis of the third most common diagnosis – psychosis – can be found. About four out of ten internees (43.9%, *n* = 233) were diagnosed with psychosis. Among the psychotic disorders, most frequent diagnoses were: schizophrenic disorders (71.2%, *n* = 166), psychoses not otherwise specified (NOS) (15.5%, *n* = 36) and delusional disorders (8.2%, *n* = 19). Psychotic disorders were mostly found in comorbidity (76.4%, *n* = 178).

Degrees of intellectual deficits, the fourth most common diagnosis (23%, *n* = 122) are shown in Table 8. Borderline intellectual functioning was found in 54 internees (44.3%); mild intellectual disability was found in 55 internees (45.1%) and moderate intellectual disability in 13 internees (10.7%). None of the internees had a severe intellectual disability. With one exception, all of the intellectual deficits were comorbid disorders. In group with borderline intellectual functioning 72.5% (*n* = 37/51[5]) followed special education or had no education. In the intellectually disabled group[6] 86.2% had special education (49/58) or no education (1/58), while

13.8% had regular education (*n* = 8). In case regular education was followed, one internee was only in kindergarten and four internees only had primary school.

Table 7. Psychotic disorders

		Total population		Population with psychotic disorder	
	N	*n*	%	*n*	%
Psychotic disorder	531	233	43.9	233	100
Schizophrenia		166	31.3	166/233	71.2
Psychosis NOS		36	6.8	36/233	15.5
Delusional disorder		19	3.6	19/233	8.2
Schizoaffective disorder		8	1.5	8/233	3.4
Psychosis due to somatic illness		2	0.4	2/233	0.9
Schizophreniform disorder		1	0.2	1/233	0.4
Brief psychotic disorder		1	0.2	1/233	0.4
Comorbidity				233	100
Sole diagnosis		55		55/233	23.6
Combi Axis I SUD and II		78		78/233	33.5
Combi Axis I SUD or II		100		100/233	42.9
Combi Axis I SUD		37			
Combi Axis II		63			

Note. Psychosis NOS = psychosis not otherwise specified.

Table 8. Intellectual deficiency

		Total population		Population with ID	
	N	*n*	%	*n*	%
Intellectual deficiency	531	122	23.0	122	100
Borderline intellectual functioning		54	10.2	54/122	44.3
Mild intellectual disability		55	10.4	55/122	45.1
Moderate intellectual disability		13	2.4	13/122	10.7
Comorbidity					
Sole diagnosis		1	0.2	1/122	0.8
Combi Axis I and Axis II		73	13.7	73/122	59.8
Combi Axis I or Axis II		48	9.0	48/122	39.3
Combi Axis I		37			
Combi Axis II		11			

Note. ID = intellectual deficiency.

Impulse control disorders were present in 14.5% (*n* = 77) of the population. All other diagnoses were present in less than 10% of the population (Table 9).

Table 9. Other diagnoses prevalent in less than 15% of the population

		Total population	
	N	*n*	%
Impulse control disorders	531	77	14.5
Anxiety and mood disorders		35	6.6
Paraphilia		23	4.3
Conduct disorders		17	3.2
Pervasive developmental disorders		16	3.0
ADHD		15	2.8
Personality disorder due to somatic cause		9	1.7
Cognitive disorders (dementia)		6	1.1
Partner relational problems		6	1.1
Adjustment disorders		3	0.6
Somatoform disorders		2	0.4
Factitious disorders		2	0.4
Learning disorders		1	0.2
Antisocial behavior		1	0.2
Psychiatric illness due to somatic illness		1	0.2
Eating disorders		1	0.2

Note. ADHD = attention deficit hyperactivity disorder.

DSM-IV-TR primary diagnoses

Besides the cumulative diagnoses, Axis I diagnoses are presented according to the primary diagnoses in Table 10. It should be noted that these primary diagnoses were based on the clinical appreciation of the author because no primary diagnoses were provided by the treating psychiatrists. In 93.4% of the cases (*n* = 496), one or more Axis I diagnoses was found. When only the primary diagnosis was examined on Axis I, psychosis was the most frequent diagnosis (42.2%, *n* = 224). Further details on the distribution of the primary diagnosis on Axis I can be found in Table 10.

DSM-IV-TR clustered according to the MacArthur classification

DSM-IV-TR diagnoses were also clustered according to the MacArthur classification (Monahan et al., 2001) in order to get primary diagnostic categories over different axes: a) major mental disorders (MMD) (30.1%, *n* = 160), b) major mental disorders in combination with substance misuse or substance-related disorders (MMD–SUD) (28.2%, *n* = 150), c) other disorders (13.2%, *n* = 70) and d) other disorders in combination with substance-related disorders (28.4%, *n* = 151).

Table 10. Primary Axis I diagnoses

		Total population		Population with Axis I diagnosis	
	N	*n*	%	*n*	%
Major mental disorder	531	294	55.4	294	100
Psychosis		224	42.2	224/294	76.2
Anxiety/mood		27	5.1	27/294	9.2
Autism/ADHD		19	3.6	19/294	6.5
Paraphilia		13	2.4	13/294	4.4
Cognitive disorder		11	2.1	11/294	3.7
Substance misuse		147	27.7	147	100
≥ 2 substance misuse		68	12.8	68/147	46.3
Alcohol		54	10.2	54/147	36.7
Drugs		25	4.7	25/147	17.0
Other disorders		55	10.4	55	100
Impulse control disorder		26	4.9	26/55	47.3
Periodic explosive disorder		15	2.8	15/55	27.3
Conduct disorder		9	1.7	9/55	16.4
Other		5	0.9	5/55	9.1
No Axis I diagnosis		35	6.6		

Intelligence

Scores from different intelligence tests (WAIS-III: Wechsler 2005; Wechsler, 1997), older version WAIS (Wechsler 1955, 1970), Raven's Progressive Matrices (RAVEN: Raven, Raven, & Court, 1998) and the short version of the Groninger Intelligence Test (sGIT: Kooreman & Luteijn, 1987) were found in 82.1% (*n* = 436) of the population. The average IQ score was 81.6 (*SD* = 17.46, range = 45–138). Intelligence scores based on the WAIS-III were available in more than half of the population (53.5%, *n* = 284). The average IQ score based on WAIS-III was 80.5 (*SD* = 16.81, range =

48–138). There was no significant difference between the performal IQ (*M* = 80.5, *SD* = 16.41) and the verbal IQ (*M* = 81.9, *SD* = 17.46) (*p* = .05).

Risk assessment (Table 11)

Psychopathy[7]. The distribution of the PCL-R scores is presented in Figure 2. The mean PCL-R total adjusted score (*n* = 224) was 21.1 (*SD* = 6.58, range = 6–36.8), with a median score of 21.1. Mean score for Factor 1 (*n* = 215) was 8.6 (*SD* = 3.40, range 1.1–16) and for Factor 2 (*n* = 213) 10.3 (*SD* = 3.86, range 1–18). Facet scores are presented in Table 11. A third scored 25[8] or more (33.48%, *n* = 75); 8.9% (*n* = 20) scored 30 or higher.

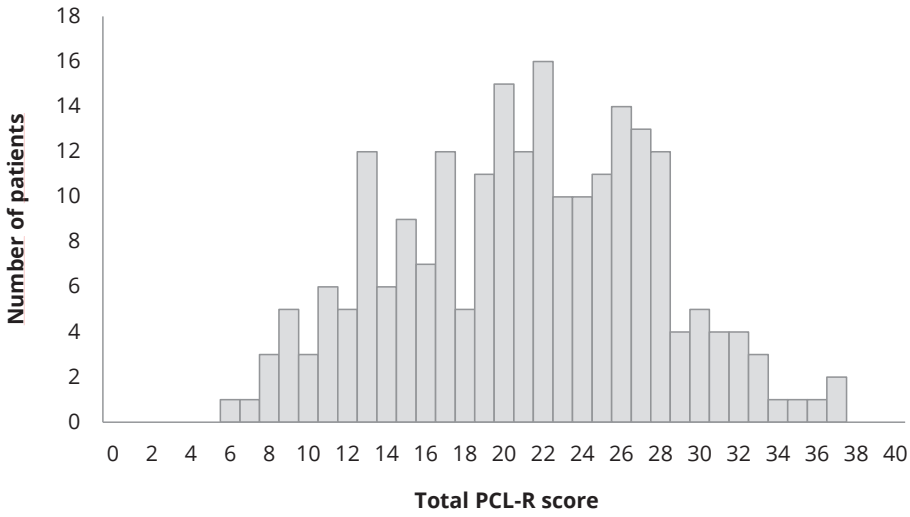


Figure 2. Distribution of PCL-R total scores.

HCR-20. The distribution of the HCR-20 scores is presented in Figure 3. The mean total HCR-20 (*n* = 298) score was 24.5 (*SD* = 5.27, range = 10.5–36.7), the mean scores for the H-, C-, and R-subscale were 14.1 (*SD* = 3.22, range = 5–20), 4.7 (*SD* = 1.83, range = 0–9) and 5.8 (*SD* = 2.12, range = 0–10), respectively. About a third (38.4%, *n* = 101) of the patients who were assessed with the HCR-20 were classified as high risk, 47.5% as medium risk (*n* = 125) and 14.1% as low risk (*n* = 37).

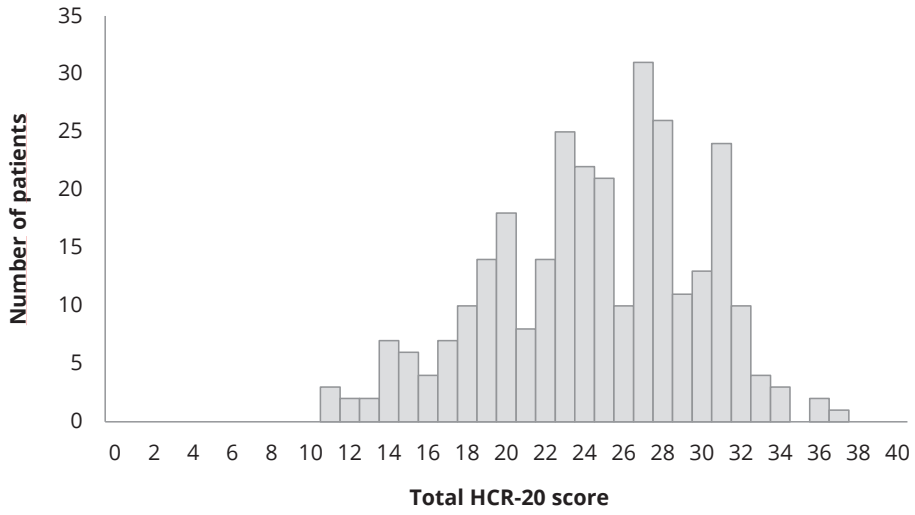


Figure 3. Distribution of HCR-20 scores.

VRAG. The distribution of the VRAG scores is presented in Figure 4. The mean score on the VRAG (*n* = 230) was 8 (*SD* = 10.94, range = -26 to +38). Further, 30.4% (*n* = 70) was categorized as high risk offender according to the three highest bins of the VRAG.

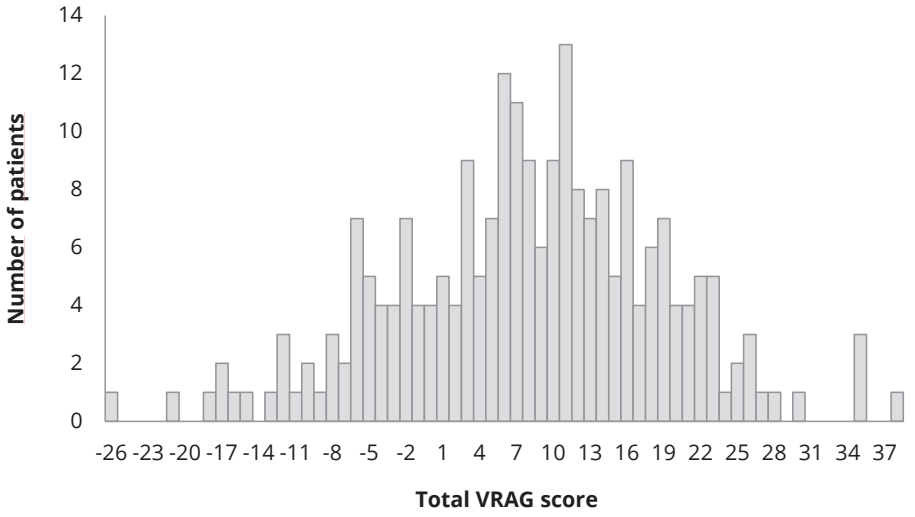


Figure 4. Distribution of VRAG scores.

Taken together, the prototypical medium security patient is diagnosed with multiple problems. Frequency analyses showed that the modus was a psychotic patient with a comorbid personality disorder and substance misuse. A third of the population was classified as being at high risk of (violent) reoffending according to different risk assessment tools.

Table 11. Risk assessment scores

		Total population			
	N	*n*	%	*M*	*SD*
HCR-20	298			24.5	5.27
H-scale	298			14.1	3.22
C-scale	298			4.7	1.83
R-scale	298			5.8	2.12
SPJ high	263	101	38.4		
PCL-R	224			21.2	6.58
Factor 1	215			8.6	3.40
Factor 2	213			10.3	3.86
Facet 1	172			3.1	2.13
Facet 2	173			5.6	1.90
Facet 3	170			6.1	2.53
Facet 4	163			5.5	2.56
Score ≥ 25	224	75	33.5		
Score ≥ 30	224	20	8.9		
VRAG	230				
VRAG score high (bin 7–9)	230	70	30.4	8	10.94

Note. HCR-20 = Historical, Clinical, Risk management-20; SPJ = structured professional judgment; PCL-R = Psychopathy Checklist-Revised; VRAG = Violence Risk Appraisal Guide.

Treatment characteristics

According to the hospital where the first MSU admission took place, 28.8% of the population was treated in Bierbeek (*n* = 153), 30.1% in Rekem (*n* = 160), and 41.1% in Zelzate (*n* = 218). Forty-eight patients were treated in two MSUs and three in all three MSUs, but only seven patients were transferred directly between MSUs. On average, the population was interned for 1,603.2 days before admission (*SD* = 1,915.7, range = 1.0–12,018.0). Most (93.8%, *n* = 498) were admitted to a MSU from a detention situation. This detention period lasted on average (*n* = 497[9]) 896.7 days (*SD*

= 1,100.4, range = 6.0–12,074.0). More than half of the population (52.9%, *n* = 281) was admitted to a MSU more than once. On average, the 281 internees with multiple admissions were readmitted 2.3 times[10] (*SD* = 2.00, range = 1–17). One hundred twenty-eight patients were readmitted once, 103 were admitted two or three times, and 50 were admitted more than four times. The distribution of the length of medium security stay until census date per category is presented in Figure 5. The average length of stay in these units was 676.4 days (*SD* = 505.9; range = 8–2,729).[11]

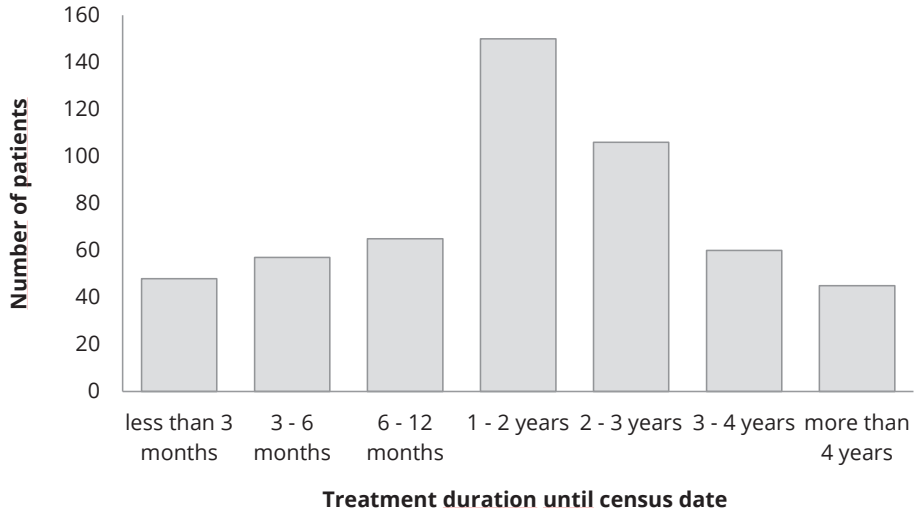


Figure 5. Distribution of treatment duration until census date.

Six patients died during medium security treatment. During the total study period, 18 patients died, nine due to suicide and one due to accidental drowning. Four of the suicides took place during medium security treatment, two in detention, one in a regular psychiatric unit, and two in the community. Twenty-four suicide attempts were registered in 18 internees.

About 10% (*n* = 65) was still in treatment at the census date. Five patients died during admission. Of the remaining group (*n* = 461), one-third (31.7%, *n* = 146) dropped out from treatment and two-thirds completed treatment (68.3%, *n* = 315). Finally, we investigated the

patients' destination at first discharge and for all discharges (Table 12). When all discharges – first admission and re-admissions – were considered (*n* = 1,123), the security levels were as follows: high security (25.2%, *n* = 283), medium security (1.6%, *n* = 18), low security (42.7%, *n* = 479), community care (24.7%, *n* = 277), and no security (5.9%, *n* = 66). Discharge to prison (high security) occurred either because of breach of conditions or new offenses. Discharge to another MSU occurred rarely. In most cases (42.7%, *n* = 479), the treatment continued in a low-security unit; of these, mostly forCare units (84.8%, *n* = 406), and a minority general psychiatric units (15.2%, *n* = 73). A quarter of the population was discharged to the community: to a forensic sheltered home (56.0%, *n* = 155), a sheltered home (14.8%, *n* = 41), or to live alone (29.2%, *n* = 81). In the case of community discharge, follow-up was provided for by the forensic team from the hospital (69.7%, *n* = 193) or a general team (30.3%, *n* = 106). Finally, a minority was discharged without security (*n* = 66), either because of death during MSU treatment (9.1%, *n* = 6) or absconding (90.9%, *n* = 60).

Table 12. Location and security level at discharge(s)

	First discharge (*n* = 507)		All discharges (*n* = 1,123)	
	n	%	*n*	%
High security (prison)	160	31.6	283	25.2
Medium security	7	1.4	18	1.6
Low security	183	36.1	479	42.7
Forensic psychiatry (forCare)	142	28.0	406	36.2
General psychiatry	41	8.1	73	6.5
Community	119	23.5	277	24.7
Forensic sheltered home	52	10.3	155	13.8
General sheltered home	21	4.1	41	3.7
Living alone	46	9.1	81	7.2
No security	38	7.5	66	5.9
Absconded	33	6.5	60	5.3
Death	5	1.0	6	0.5

DISCUSSION

After an offender has been found NGRI in Belgium, one of the first goals is to assess the level of security needed to ensure the safety to the public, the staff, and the patients of the institution. The systematic allocation of patients to appropriate security levels is central to the operation of forensic mental health services. However, in the UK, MSU use is suboptimal (Coid & Kathan, 2000), with patients on the waiting list outnumbering the available beds by a factor of five-to-one (Melzer et al., 2004). On the one hand, a shortage of long-term medium security provisions and hence admissions of patients requiring long-term care were noted. On the other hand, patients who did not need medium security treatment were admitted anyways because local services could not cope with them (Melzer et al., 2004), or because a parallel circuit of readmissions emerged in ex-patients followed in the community by forensic services (Coid & Kathan, 2000). In Flanders, indications for referral to a MSU were made clinically, which raises the question of just how reliable, valid, and transparent assessments are in a real-world setting (Stredny, Parker, & Dibble, 2012). Describing if and how a population matches a medium security level is not an easy task. Besides basic considerations such as the height of the fence or staffing levels, other factors become far more difficult to quantify. Also, there is an inherent problem of circularity in seeking to define the patient characteristics appropriate for a type of secure unit or level of security if one describes those characteristics by patients currently in such units. In this respect, there remains a need to measure the elements of different levels of security within forensic psychiatric services more clearly. Notwithstanding, in this chapter, attempts were made to outline the characteristics of the study population, which were give the label medium security on a clinical basis.

The present study confirmed that admitted patients generally had long criminal and psychiatric histories and poor social networks. In sum, the following profile of a medium security internee emerged:

- Belgian male between 25 and 34 years old;
- diagnosed with a combination of psychosis, substance misuse, and personality disorder;
- previously admitted to a general psychiatric hospital;
- without high school diploma and with an unstable work and relationship history;
- living alone, having no partner or children at the time of the index offense;

- interned for a violent index offense; and
- on average 24.8 years old when first sentenced, with an average of 6.2 prior sentences and 4.6 years of detention prior to first admission in medium security.

Comparing forensic psychiatric populations across countries should be interpreted with caution because of major differences between the legal systems and related organization of forensic health care, as well as variations in the characteristics of local patient groups and local treatment providers (Melzer et al., 2004; Salize & Dressing, 2007). Notwithstanding these and other obstacles discussed in the introduction, several conclusions can be made.

Regarding sociodemographic characteristics, few differences with the literature were noted. Age, nationality, education, as well as relationship and work status were all quite similar (Blattner & Dolan, 2009; Coid et al., 2001; Dolan & Khawaja, 2004; Gow et al., 2010; Melzer et al., 2004). In comparison with internees in Wallonia (41.6 years), however, the mean age of Flemish internees (36.5 years) was somewhat lower. Discharge to the community was less frequent in comparison to other studies (Blattner & Dolan, 2009; Dolan & Khawaja, 2004; Gow et al., 2010). Readmissions to medium security were higher than or comparable with other research studies (22%, Coid et al., 2001; 47%, Dolan & Khawaja, 2004). In line with the literature (Blattner & Dolan, 2009; Gow et al., 2010; Linhorst & Scott, 2004; Melzer et al., 2004), the high percentage of internees with prior admissions to general psychiatric services was striking, and calls into question the role of general mental health services. For example Brand, Mellsop, and Tapsell (2015) examined psychiatric care provided in the year prior to offending and found that access to care was not the problem, whereas a non-assertive approach to treatment was. Non-compliance with general psychiatric care was associated with being assessed as needing medium security care (Melzer et al., 2004). However, whether there is a causal relationship between effective psychiatric care in the sense of symptom reduction and offense prevention is less clearly established (Brand et al., 2015).

In terms of clinical diagnoses, high rates of comorbidity were found, in line with other studies (Blattner & Dolan, 2009; Gow et al., 2010). Internees were less likely to have a psychotic disorder and more likely to have a personality disorder than international samples (Blattner & Dolan, 2009; Coid et al., 2001; Dolan & Khawaja, 2004; Gow et al., 2010; Lelliott et al., 2001; Melzer

et al., 2004). In French-speaking internees, the same number of personality disordered internees (71%) was found by Pham and Saloppé (2010). The small number of psychotic internees (18%) found in the same study can be explained by the inclusion of stabilized patients only (Pham & Saloppé, 2010). Substance misuse comorbidity was comparable to or higher than rates found in international studies (Blattner & Dolan, 2009; Carr et al., 2006; Dolan & Khawaja, 2004; Gow et al., 2010; Gradillas et al., 2007; Hildebrand & de Ruiter, 2004) and lower than findings in the south of the country (Pham & Saloppé, 2010). Cluster B, more specifically APDs, were the most common personality disorders, which was in line with other research (Dolan & Khawaja, 2004; Hildebrand & de Ruiter, 2004; Pham & Saloppé, 2010; Timmerman & Emmelkamp, 2001).

With regard to judicial history, few first offenders were found and index offenses were mainly violent offenses. While these findings are in line with the literature (Blattner & Dolan, 2009; Dolan & Khawaja, 2004; Freestone et al., 2012; Gow et al., 2010; Lelliott et al., 2001; Melzer et al., 2004), the number of prior convictions in the study population was rather high (e.g., 6.2 vs. 3.5; Freestone et al., 2012). In addition, it is interesting to note that 80% of internees had previously been convicted and one-third had prior internments (i.e., different assessments regarding accountability were regularly made). This is perhaps not surprisingly, since previous research revealed that a significant proportion of the male and female population residing in prisons suffer from psychotic, mood, personality, and substance use disorders (Diamond, Wang, Holzer, Thomas, & des Anges, 2001; Fazel & Seewald, 2012). As Rice and Harris (1997) noted, people tend to shift back and forth between the mental health and criminal justice systems.

Mean scores on risk assessment instruments in the study population were somewhat higher than the scores found in medium security samples in the UK, differing by about five points on the total HCR-20 scale, three points on the H-scale, one point on the C-scale and one point on the R-scale (Dolan & Khawaja, 2004; Gray et al., 2004; Gray et al., 2008). Similar scores were found in a medium to high risk sample in the French-speaking part of Belgium (Wallonia; Claix & Pham, 2004) and in forensic psychiatric hospitals in the Netherlands (de Vogel & de Ruiter, 2006; Mudde, Nijman, van der Hulst, & van den Bout, 2011). The mean VRAG score in the current study was somewhat higher than the construction sample described in the manual (Quinsey et al., 2006), but comparable to other studies (Gray et al., 2007; Ho et al., 2009; Snowden et al., 2010; Snowden et al., 2009; Snowden et al., 2007). The average total PCL-R score was similar to the mean score

for forensic psychiatric patients reported by Hare (2003), and slightly higher than those found by Pham and Saloppé (2010).

Surprisingly, risk assessment instruments and intelligence scores were not assessed on a regular basis, which makes it difficult to construct patient profiles (van der Veeken, Bogaerts, & Lucieer, 2015). Also, assessing risk is considered an important aspect when following the RNR-model.

In terms of treatment characteristics, there were significant dropout rates. It could be hypothesized that the population at hand was not always adequately matched to the medium security level that was offered, with at least a subpopulation probably needing a higher security level. During the study period, there were no high security beds in Flanders, and, as argued by Coid and Kathan (2000), the absence of alternatives can play a role in admission criteria. Another hypothesis is that treatment programs were not responsive enough to offender characteristics. For instance, a large number of intellectually disabled patients were identified, whereas the MSUs in principle do not provide treatment for this particular group. From the patient's perspective, two previous studies on interned offenders revealed that internees had more positive and less negative experiences in prison settings when compared to care settings (De Smet et al., 2015; To, Vanheule, De Smet, & Vandevelde, 2015). Internees in treatment settings reported a lack of control and experienced too much pressure to perform in treatment and a constant fear of being sent back to prison, as compared to internees in prison settings (To et al., 2015). Further, prison staff was perceived as more genuinely investing in their well-being, whereas the authenticity of treatment staff was questioned (De Smet et al., 2015; To et al., 2015). The studies highlighted the importance of fulfilling (therapeutic) relationships with their therapists and the need for being in charge of personal choices (De Smet et al., 2015). Given the fact that these two studies were small exploratory studies with substantial selection bias, further research in this area is recommended.

Limitations and strengths

First of all, this multi-center, retrospective study was based on case note material obtained for clinical rather than for research purposes. Not only were there several hospitals involved, there were also several clinicians, and recording systems differed in rigor. This resulted in missing data, which were very evident when examining the risk assessment instruments. Also, inconsistencies

were found regarding the MPD diagnoses, which were adjusted in almost half of all cases. In the UK, Gow et al. (2010) noted that a review of case notes also suggested much higher rates of substance misuse than reported. Regarding judicial data, De Vuysere et al. (2004) noted in her research project examining CPS files of three Flemish CPSs that there was no uniformity in CPS files. For example, a Central Criminal Record and a primary psychiatric diagnosis were not consistently found, and certain characteristics had no standardized notation. The present study was in line with these findings. It is important to note, however, that a genuine interest in research and cooperation was encountered during the CPS data collection, as was also the case with previous research (De Vuysere et al., 2004).

While the limitations of the study are important, so are its strengths. The study virtually covered all medium security admissions over an extensive period in Flanders. The data collection was undertaken very carefully, and most data were cross-referenced. Furthermore, the study enabled research on the field validity of risk assessment instruments, a neglected area in the literature.

CONCLUSION

This chapter provided an in-depth analysis of the characteristics of the medium security population in Flanders. Whether or not the security level totally matched the population at hand is a difficult question to answer. However, the study revealed a number of arguments in favor of using categorical medium security care, since identified characteristics were very similar to medium to high security populations internationally. In the future, it is recommended to use instruments to measure security level more in detail. Furthermore, it would be interesting to identify distinct profiles of medium security subpopulations and investigate if the medium security approach *works* better for specific subpopulations than for others. For example, van der Veeken et al. (2015) identified four patient profiles with varying psychopathologies, risk factors, and crimes, which could enhance treatment guidelines. MSU beds are scarce and expensive and should be used optimally. As was noted in the UK, this was not always the case; almost one out of four patients admitted to medium security did not actually need this level of security (Coid & Kathan, 2000).

FOOTNOTES

[1] The study makes no specifications of internees having both correctional and internment priors.

[2] Valid percentage based on IQ below 70.

[3] Valid percentage based on IQ below 75.

[4] Regular, technical or vocational education.

[5] Three missings.

[6] Ten missings.

[7] Although the PCL-R is a diagnostic, rather than a risk assessment instrument, psychopathy being an important risk factor is included in this section.

[8] Scores of 25 and higher on the PCL-R are indicative of psychopathy in Europe (Cooke & Michie, 1999).

[9] One missing in SIDIS.

[10] Including readmission in another medium security unit.

[11] For each internee the duration of stay was calculated by combining first admission and possible readmissions. For the patients still residing in a medium security unit on 31/12/2010, admission duration was calculated to that date.

REFERENCES

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17*, 19-52.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Barrett, S. P., Plooy, J. du, Toit, J. du, Wilmans, S., Calitz, F. J. W., & Joubert, G. (2007). Profile of mentally ill offenders referred to the Free State Psychiatric Complex, 2004 - 2008. *South African Journal of Psychiatry, 13*, 56-58.

Blattner, R., & Dolan, M. (2009). Outcome of high security patients admitted to a medium secure unit: The Edenfield Centre study. *Medicine, Science, and the Law, 49*(4), 247-256. doi: 10.1258/rsmmsl.49.4.247

Boers, A., Vandevelde, S., Soyez, V., Smet, S. De, & To, W. T. (2011). Het zorgaanbod voor geïnterneerden in België. *Panopticon, 32*(2), 17-38.

Brand, E. (2005). *Onderzoeksrapport PIJ-dossier 2003-C. Predictieve validiteit van de FPJ-lijst*. Den Haag: DJI.

Brand, E., Mellsop, G., & Tapsell, R. (2015). A review of the psychiatric care provided to patients who subsequently offended. *International Journal of Clinical Psychiatry and Mental Health, 3*, 31-35. doi: 10.12970/2310-8231.2015.03.01.5

Carr, W. A., Rotter, M., Steinbacher, M., Green, D., Dole, T., Garcia-Mansilla, A., . . . Rosenfeld, B. (2006). Structured Assessment of Correctional Adaptation (SACA): A measure of the impact of incarceration on the mentally ill in a therapeutic setting. *International Journal of Offender Therapy and Comparative Criminology, 50*(5), 570-581. doi: 10.1177/0306624X06289176

Casselman, J. (2011). Internering in Vlaanderen anno 2010: Ondanks het goede nieuws nog steeds mensonwaardig. In I. Aertsen, T. Daems, A. Haaren, F. Hutsebaut, & J. Maesschalck (Eds.), *Deviante wetenschap. Het domein van de criminologie. Liber amicorum Johan Goethals* (pp. 237-244). Leuven: Acco.

Casselman, J., Vuysere, S. De, & Vervaeke, G. (2003a). De Commissies ter Bescherming van de Maatschappij van Antwerpen, Gent en Leuven. Een vergelijking van de geïnterneerdenpopulatie. *Panopticon, 6*, 604-609.

Casselman, J., Vvuysere, S. De, & Vervaeke, G. (2003b). De geïnterneerdenpopulatie van Vlaanderen: Enkele cijfergegevens. *Panopticon, 6*, 599-604.

Casselman, J., Pelc, I., Servais, L., Heylen, A., Sintobin, M., Harq, E., . . . Hoffman, D. (2002). *Evaluatie van Forensisch Psychiatrische Eenheden. Pilootproject: Het aanbod van een intensieve klinische behandeling voor de geïnterneerden met als doel een zo groot mogelijke resocialisatie van die patiënten*: Federaal Ministerie van Sociale Zaken, Volksgezondheid en Leefmilieu.

Claix, A., & Pham, T. H. (2004). Evaluation of the HCR-20 Violence Risk Assessment Scheme in a Belgian forensic population. *L'Encéphale, 30*(5), 447-453.

Coid, J., Kahtan, N., Gault, S., Cook, A., & Jarman, B. (2001). Medium secure forensic psychiatry services: Comparison of seven English health regions. *British Journal of Psychiatry, 178*(1), 55-61.

Coid, J., & Kathan, N. (2000). An instrument to measure the security needs of patients in medium security. *The Journal of Forensic Psychiatry, 11*(1), 119-134.

Collins, M., & Davies, S. (2005). The security needs assessment profile: A multidimensional approach to measuring security needs. *International Journal of Forensic Mental Health, 4*(1), 39-52. doi: 10.1080/14999013.2005.10471211

Cooke, D. J., & Michie, C. (1999). Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology, 108*(1), 58–68. doi: 10.1037/0021-843X.108.1.58

Cosyns, P. (2005). *Werkgroep forensisch psychiatrisch zorgcircuit: Syntheseverslag mei 2005*. Brussel: Federale Overheidsdienst Justitie.

Cosyns, P. (2005, September). *Behandeling van geïnterneerden in Vlaanderen: Een uitdaging voor de toekomst*: Paper presented at the Herfstvergadering VVP, Grimbergen, Belgium.

Cosyns, P., D'Hont, C., Janssens, D., Maes, E., & Verellen, R. (2007). Geïnterneerden in België. De cijfers. *Panopticon, 1*, 46-61.

Cosyns, P., Dillen, C., Ruyter, B. de, & Doncker, D. De (1994). Behandelingsmogelijkheden van geinterneerden in algemeen psychiatrische ziekenhuizen. In F. Lampaert (Ed.), *Gevangenis en samenleving II* (pp. 75-118). Brussel: Koning Boudewijnstichting.

Deckers, A., Seynnaeve, K., Smedt, A. De, Dheedene, J., Vandenplas, E., & Auwera, A. Van der (2014). *Geïnterneerden in detentie op 24/12/2013*. Unpublished report.

Diamond, P. M., Wang, E. W., Holzer, C. E., 3rd, Thomas, C., & des Anges, Cruser. (2001). The prevalence of mental illness in prison. *Administration and Policy in Mental Health, 29*(1), 21-40.

Dolan, M., & Khawaja, A. (2004). The HCR–20 and post-discharge outcome in male patients discharged from medium security in the UK. *Aggressive Behavior, 30*(6), 469-483. doi: 10.1002/ab.20044

Doncker, D. De, Koeck, S., Huys, W., & Winter, J. (2006). Classificatie van extrafamiliale pedoseksuelen, Deel I: Het MTC:CM3 systeem toegepast binnen een ambulante setting. *Tijdschrift voor Seksuologie, 30*(1), 16-26.

Fazel, S., & Seewald, K. (2012). Severe mental illness in 33,588 prisoners worldwide: Systematic review and meta-regression analysis. *British Journal of Psychiatry, 200*(5), 364-373. doi: 10.1192/bjp.bp.111.096370

Flynn, G., O'Neill, C., McInerney, C., & Kennedy, H. G. (2011). The DUNDRUM-1 structured professional judgment for triage to appropriate levels of therapeutic security: Retrospective-cohort validation study. *BMC Psychiatry, 11*, 43. doi: 10.1186/1471-244X-11-43

Freestone, M., Taylor, C., Milsom, S., Mikton, C., Ullrich, S., Phillips, O., & Coid, J. (2012). Assessments and admissions during the first 6 years of a UK medium secure DSPD service. *Criminal Behaviour and Mental Health, 22*(2), 91-107. doi: 10.1002/cbm.1823

Goethals, J. (1985). Internering in België. Enkele statische gegevens. *Panopticon, 6*(2), 129-147.

Gow, R. L., Choo, M., Darjee, R., Gould, S., & Steele, J. (2010). A demographic study of the Orchard Clinic: Scotland's first medium secure unit. *The Journal of Forensic Psychiatry & Psychology, 21*(1), 139-155. doi: 10.1080/14789940903188956

Gradillas, V., Williams, A., Walsh, E., & Fahy, T. (2007). Do forensic psychiatric inpatient units pose a risk to local communities? *The Journal of Forensic Psychiatry & Psychology, 18*(2), 261-265. doi: 10.1080/14789940701246360

Gray, N. S., Fitzgerald, S., Taylor, J., MacCulloch, M. J., & Snowden, R. J. (2007). Predicting future reconviction in offenders with intellectual disabilities: The predictive efficacy of VRAG, PCL–SV, and the HCR–20. *Psychological Assessment, 19*(4), 474–479. doi: 10.1037/1040-3590.19.4.474

Gray, N. S., Snowden, R. J., MacCulloch, S., Phillips, H., Taylor, J., & MacCulloch, M. J. (2004). Relative efficacy of criminological, clinical, and personality measures of future risk of offending in mentally disordered offenders: A comparative study of HCR-20, PCL:SV, and OGRS. *Journal of Consulting and Clinical Psychology, 72*(3), 523-530. doi: 10.1037/0022-006X.72.3.523

Gray, N. S., Taylor, J., & Snowden, R. J. (2008). Predicting violent reconvictions using the HCR-20. *The British Journal of Psychiatry, 192*(5), 384-387. doi: 10.1192/bjp.bp.107.044065

Habets, P., Jeandarme, I., Uzieblo, K., Oei, K., & Bogaerts, S. (2014). Intelligence is in the eye of the beholder: Investigating repeated IQ measurements in forensic psychiatry. *Journal of Applied Research in Intellectual Disabilities, 28*(3), 182-192. doi: 10.1111/jar.12120

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist*. Toronto, ON: Multi-Health Systems.

Hende, M. Vanden, Caris, K., & Block-Bury, L. De (2005). *Ontgrendeld. Beschrijvend wetenschappelijk onderzoek naar geïnterneerden met een verstandelijke handicap en hun verblijf in Vlaamse gevangenissen*. Gent: Academia Press.

Hildebrand, M., & Ruiter, C. de (2004). PCL-R psychopathy and its relation to DSM-IV Axis I and II disorders in a sample of male forensic psychiatric patients in The Netherlands. *International Journal of Law and Psychiatry, 27*(3), 233-248. doi: 10.1016/j.ijlp.2004.03.005

Ho, H., Thomson, L., & Darjee, R. (2009). Violence risk assessment: The use of the PCL-SV, HCR-20, and VRAG to predict violence in mentally disordered offenders discharged from a medium secure unit in

Scotland. *The Journal of Forensic Psychiatry & Psychology, 20*(4), 523-541. doi: 10.1080/14789940802638358

Ibishi, N. F., Musliu, N. R., Hundozi, Z., & Citaku, K. (2015). Prevalence of the hospitalisation of mentally ill offenders in the Forensic Unit of the Clinic of Psychiatry in Pristina over a three-year period and long-term strategy implications for the management of the Forensic Mental Health System Service. *Medicine, Science, and the Law, 55*(3), 201-207. doi: 10.1177/0025802414532247

IBM Corp. (Released 2013). *IBM SPSS Statistics for Windows, Version 22.0*. Armonk, NY: IBM Corp.

Kennedy, H. G. (2002). Therapeutic uses of security: Mapping forensic mental health services by stratifying risk. *Advances in Psychiatric Treatment, 8*, 433-443. doi: 10.1192/apt.8.6.433

Knight, R., Carter, D., & Prentky, R. (1989). A system for the classification of child molesters: Reliability and application. *Journal of Interpersonal Violence, 4*(1), 3-23.

Kooreman, A., & Luteijn, F. (1987). *The short format of the Groningen Intelligence Test*. Lisse: Swets & Zeitlinger.

Lelliott, P., Audini, B., & Duffett, R. (2001). Survey of patients from an inner-London health authority in medium secure psychiatric care. *British Journal of Psychiatry, 178*(1), 62-66. doi: 10.1192/bjp.178.1.62

Linhorst, D., & Scott, L. P. (2004). Assaultive behavior in state psychiatric hospitals. Differences between forensic and nonforensic patients. *Journal of Interpersonal Violence, 19*(8), 857-874. doi: 10.1177/0886260504266883

Maes, B., Goethals, J., & Verlinden, S. (2009). *Personen met een verstandelijke handicap onderhevig aan een interneringsmaatregel*. Leuven: Steunpunt Welzijn, Volksgezondheid en Gezin.

Melzer, D., Tom, B., Brugha, T., Fryers, T., Gatward, R., Grounds, A., . . . Meltzer, H. (2004). Access to medium secure psychiatric care in England and Wales. 1: A national survey of admission assessments. *The Journal of Forensic Psychiatry & Psychology, 15*(1), 7-31. doi: 10.1080/14789940410001661856

Menezes, S., Oyebode, F., & Haque, M. (2009). Victims of mentally disordered offenders in Zimbabwe, 1980–1990. *The Journal of Forensic Psychiatry & Psychology, 20*(3), 427-439. doi: 10.1080/14789940802542907

Monahan, J., Steadman, H. J., Silver, E., Appelbaum, P., Robbins, P., Mulvey, E., . . . Banks, S. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York, NY: Oxford University Press.

Mudde, N., Nijman, H., Hulst, W. van der, & Bout, J. van den (2011). Het voorspellen van agressie tijdens de behandeling van forensisch psychiatrische patiënten aan de hand van de HCR-20 [Predicting aggression during the treatment of forensic psychiatric patients by means of the HCR-20]. *Tijdschrift voor Psychiatrie, 53*(10), 705-713.

Pal, S. (1997). Mental disorders in abnormal offenders in Papua New Guinea. *Medicine and Law, 16*(1), 87-95.

Pham, T., Saloppé, X, Bongaerts, X., & Hoebanx, J.L. (2007). L'expertise dans le cadre de la loi de Défense Sociale en Belgique: Repères diagnostiques et recommandations. *Annales Médico Psychologiques, 165*, 49–55.

Pham, T., & Saloppé, X. (2010). PCL-R Psychopathy and its relation to DSM Axis I and II disoders in a sample of male forensic patients in a Belgian security hospital. *International Journal of Forensic Mental Health, 9*(3), 205 — 214. doi: 10.1080/14999013.2010.517255

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Raven, J., Raven, J.C., & Court, J.H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* San Antonio, TX: Harcourt Assessment.

Rice, M., & Harris, G. (1997). The treatment of mentally disordered offenders. *Psychology, Public Policy, and Law, 3*(1), 126-183.

Rycke, R. De, & Pauwelyn, L. (2015). Zorg voor geïnterneerde personen: Algemene organisatie en beleidsaanbevelingen. In J. Casselman, R. De Rycke, & H. Heimans (Eds.), *Internering* (pp. 179-208). Brugge: die Keure.

Salize, H. J., & Dressing, H. (2007). Admission of mentally disordered offenders to specialized forensic care in fifteen European Union member states. *Social Psychiatry and Psychiatric Epidemiology, 42*(4), 336-342. doi: 10.1007/s00127-007-0159-2

Schuringa, E., Spreen, M., & Bogaerts, S. (2014). Inter-rater and test-retest reliability, internal consistency, and factorial structure of the instrument for forensic treatment evaluation. *Journal of Forensic Psychology Practice, 14*(2), 127-144. doi: 10.1080/15228932.2014.897536

Smedt, A. De, Mariën, L., & Vermeiren, R. (2008, September). *Profilering van de medium risk geïnterneerden in forensische psychiatrische settings.* Paper presented at the Fourth Vlaams Geestelijk Gezondheidscongres, Ghent, Belgium.

Smet, S. De, Hecke, N. Van, Verte, D., Broekaert, E., Ryan, D., & Vandevelde, S. (2015). Treatment and control: A qualitative study of older mentally ill offenders' perceptions on their detention and care trajectory. *International Journal of Offender Therapy and Comparative Criminology, 59*(9), 964-985. doi: 10.1177/0306624X14521129

Snowden, R. J., Gray, N. S., & Taylor, J. (2010). Risk assessment for future violence in individuals from an ethnic minority group. *International Journal of Forensic Mental Health, 9*(2), 118-123. doi: 10.1080/14999013.2010.501845

Snowden, R. J., Gray, N. S., Taylor, J., & Fitzgerald, S. (2009). Assessing risk of future violence among forensic psychiatric inpatients with the Classification of Violence Risk (COVR). *Psychiatric Services, 60*(11), 1522-1526. doi: 10.1176/appi.ps.60.11.1522

Snowden, R. J., Gray, N. S., Taylor, J., & MacCulloch, M. J. (2007). Actuarial prediction of violent recidivism in mentally disordered offenders. *Psychological Medicine, 37*(11), 1539-1549. doi: 10.1017/s0033291707000876

Spreen, M., Brand, E., Horst, P. Ter, & Bogaerts, S. (2014). *Handleiding en methodologische verantwoording HKT-R, Historisch, Klinische en Toekomstige-Revisie*. Groningen: Dr. van Mesdag.

Stredny, R. V., Parker, A. L., & Dibble, A. E. (2012). Evaluator agreement in placement recommendations for insanity acquittees. *Behavioral Sciences & the Law, 30*(3), 297-307. doi: 10.1002/bsl.1995

Strydom, N., Pienaar, C., Dreyer, A., Merwe, L. van der, & Jansen van Rensburg, B. (2011). Profile of forensic psychiatric inpatients referred to the Free State Psychiatric Complex, 2004 - 2008. *South African Journal of Psychiatry, 17*, 40-43.

Timmerman, I., & Emmelkamp, P. (2001). The Prevalence and Comorbidity of Axis I and Axis II Pathology in a Group of Forensic Patients. *International Journal of Offender Therapy and Comparative Criminology, 45*(2), 198-213.

To, W. T., Vanheule, S., Smet, S. De, & Vandevelde, S. (2015). The treatment perspectives of mentally ill offenders in medium- and high-secure forensic settings in Flanders. *International Journal of Offender Therapy and Comparatvie Criminology, 59*(14), 1605-1622. doi: 10.1177/0306624X14566355

Vandenbroucke, M. (1995). Geïnterneerde: wie bent u? Een vraag naar de diagnose van de geinterneerde in Belgie. In J. Casselman, P. Cosyns, J. Goethals, M. Vandenbroucke, D. De Doncker, & C. Dillen (Eds.), *Internering* (pp. 63-72). Leuven: Garant.

Veeken, F. C. van der, Bogaerts, S., & Lucieer, J. (2015). Patient profiles in Dutch forensic psychiatry based on risk factors, pathology, and offense. *International Journal of Offender Therapy and Comparative Criminology*, Advance online publication. doi: 10.1177/0306624X15619636

Vertommen, H., & Maesschalck, C. (2007). *Definitie en operationalisering van het concept "medium risk"*. Research report.

Vogel, V. de, & Ruiter, C. de (2006). Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime & Law, 12*(3), 321-336. doi: 10.1080/10683160600569029

Vuysere, S. De (2005). Veertig jaar wet ter bescherming van de maatschappij: Reden tot feesten? In J. Goethals, F. Hutsebaut, & G. Vervaeke (Eds.), *Gerechtelijke geestelijke gezondheidszorg: Wetenschap, beleid en praktijk. Liber amicorum Joris Casselman* (pp. 259-269). Leuven: University Press.

Vuysere, S. De, Casselman, J., & Vervaeke, G. (2004). Geïnterneerden verblijvend in een rijksinstelling in Vlaanderen. *Tijdschrift voor Welzijnswerk, 260*, 18-35.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence* (Version 2). Burnaby, BC: Simon Fraser University.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale WAIS*. New York, NY: Psychological Corporation.

Wechsler, D. (1970). *WAIS Nederlandstalige Bewerking. Technische Handleiding* (J. Stinissen, P. J. Willems, P. Coetsier & W. L. L. Hulsman, Trans.). Amsterdam: Swets & Zeitlinger.
Wechsler, D. (2005). *WAIS-III Nederlandstalige Bewerking. Technische Handleiding*. Lisse: Swets & Zeitlinger.
Wechsler, D. (1997). *The Wechsler Adult Intelligence Scale – Third Edition*. San Antonio, TX: The Psychological Corporation.
Yusuf, A. J., & Nuhu, F. T. (2009). The profile of mentally ill offenders in Katsina, Northern-Nigeria. *African Journal of Psychiatry, 12*(3), 231-232.

Abolition of the internment of convicted prisoners: Clinical implications?

Tijdschrift voor Psychiatrie, 28(8), 574-582

Jeandarme, I., Pouls, C., Hanoulle, K., Oei, T. I., & Bogaerts, S.

4

ABSTRACT

Background. According to Article 21 of the current Belgian social defense law, the Minister of Justice can transfer a convicted prisoner to the authority of the Commission of the Protection of Society when in the course of a prison sentence a severe mental illness emerges. Currently there is a proposal to abolish Article 21 and psychiatric hospitals fear to be confronted with dangerous forensic patients.

Aim. To explore and compare the profile of convicted internees to that of regular internees.

Method. Convicted internees (*n* = 48) and regular internees (*n* = 483) treated in one of the Flemish medium security units were compared on demographic, clinical and risk factors.

Results. Compared to regular internees, convicted internees showed more severe psychiatric problems as well as a higher risk profile.

Conclusion. Abolition of Article 21 poses a serious challenge to regular psychiatric hospitals.

INTRODUCTION

Under Belgian law, internment – regulated by the Act of April 9, 1930 to Protect Society (APS) from Abnormals and Habitual offenders (*Wet ter Bescherming van de Maatschappij* WBM, 1964) – is an indefinite safety measure imposed by a (investigating) judge to an offender if the latter is found not guilty by reason of insanity (NGRI). Offenders can be interned if it is proven that they have committed an offense[1] and they are found irresponsible or 'severely diminished responsible' at the moment of the trial as a consequence of either a status of insanity or a serious mental deficiency which makes the person unable to (fully) control his actions. Internment is a safety measure and not a punishment, with a dual purpose, namely the protection of society *and* the medical-psychiatric treatment of the offender (Vandevelde et al., 2011). Although somewhat similar to the measure of *Terbeschikkingstelling (TBS)* in the Netherlands, there are important differences. Internment can be imposed in case of crimes or misdemeanors with a minimum sentence of eight days, whereas TBS can only be imposed for crimes with a minimum sentence of four years or crimes out of an exhaustive list. In the Netherlands, five levels of accountability are used, which allows for combinations of punishment and safety measures, whereas in Belgium, accountability is assessed dichotomously.

In addition to regular internment measures, the Minister of Justice can impose an internment measure to a convict, when during detention the convict is found to be "in a state of insanity, a serious state of mental disturbance or mental deficiency, which renders him/her unfit to control his/her actions" (Art. 21 APS). In this so-called ministerial internment there is no need for a connection between the original criminal conviction and the mental disorder (Weis, 2010). When the mental condition is sufficiently restored before the end of the sentence, the internment is stopped. If the mental condition has not improved after the expiration of the original sentence, a ministerial circular of November 24, 1964 stipulates that the internment remains for an indefinite period (art. 24). From then on, the Commission of the Protection of Society (CPS) has jurisdiction and actually acts as a court (Van Den Berge, 2011).

Criticism of the ministerial internment

According to Dijkmans (1980), De Clerck and Van Steenbrugge (2007) and Weis (2010) fundamental basic - and human rights are violated by the ministerial internment. First, the ministerial internment is in conflict with Article 6 (*the right to a fair trial*) and Article 13 (*the right of access to court*) of the European Convention on Human Rights (ECHR). Second, the ministerial internment is in conflict with Article 12 of the Constitution and Article 5 of the ECHR (*the right to personal liberty*). The European Court formulated in the Winterwerp arrest (1979) three conditions to detain a mentally ill person in order to be compatible with Article 5 ECHR. First, the deprivation of liberty must be based on an objective medical examination showing the mental illness. Second, the mental illness must be severe and third, the deprivation of liberty may only last as long as the mental illness exists. The ECHR does not require that a judge orders the deprivation of liberty but allows an administrative authority (i.e., the Minister of Justice) to take this decision. Despite the margin of appreciation that Belgian authorities have, Weis (2010) argued that the uniform advice of the CPS, based on a brief examination by the prison psychiatrist, can hardly be equated with an objective medical expert report. Weis (2010) further argued that, strictly speaking, the law does not allow for the convicted internee to make a request for his release, although in practice the CPSs assume their authority to decide on conditional release.

In summary, in the case of a ministerial internment, a specified detention period - ordered by a judge – is changed unilaterally, on the basis of a brief examination, without trial and without contradictory process into an indefinite confinement by a ministerial decision. As such the legal status of the convicted prisoner changes into that of a convicted internee. Despite the above criticisms, the Constitutional Court (2011) found that the application of the ministerial internment did not violate the constitutional and European provisions regarding the deprivation of liberty.

Legislative change proposals

Nevertheless there was a need for a better regulation, which came with the new Act of April 21, 2007 concerning the Internment of Persons with Mental disorder (AIPM, *Wet Internering van Personen met een Geestesstoornis, WIPG,* 2007). In this Act, the internment of convicts was referred to the sentencing tribunal (*Strafuitvoeringsrechtbank, SURB*), which allowed for an adversarial process (Van Den Berge, 2008). The (extended) internment of convicts after the original detention

period was abolished and replaced by the civil law procedure on compulsory admission as provided by the Act on the Protection of Mentally ill persons (APM, *Wet betreffende Bescherming Persoon Geesteszieke, WPG*). The AIPM (2007) never came into effect and the controversial ministerial internment was completely deleted in the (first version) of the new internment act (2014). Implicitly it is assumed that a convict who develops a serious mental illness at the end of his prison term and constitutes a danger to himself and /or society, will now be compulsory admitted to a psychiatric unit under Article 22bis APM (1990).

Psychiatric care after the law reform

The question remains whether the law proposal will resolve all of the problems (Heimans, Vander Beken, & Schipaanboord, 2015). First, it would no longer be possible to transfer convicted prisoners who develop a mental disorder during incarceration to an appropriate care facility. Second, after the prison sentence has expired, these patients would be send to a regular psychiatric unit, which cannot provide or guarantee the necessary care nor the necessary security level. Such an arrangement would in other words, require a reorganization of regular psychiatric care systems. According to Heimans et al. (2015) they would be faced with mentally disordered offenders who have committed serious crimes and have severe mental illnesses, posing a real violent threat both inside and outside of an institution. Ever since the publication of the internment act in 2007, the mental health care sector was worried about this evolution. Umbrella organization *Zorgnet Vlaanderen* (as of May 5, 2015 *Zorgnet-Icuro*) emphasized that the sector assumes the engagement "to ensure the treatment of internees for each area under the Court of Appeal. The sector agrees in other words to a collective duty to admit an internee provided that the network of facilities in the area has the means to do so. The place of admission is decided by the network. An agreement between the network and the government acts as a framework" (Moens & Pauwelyn, 2012, p. 52). The National Council for Hospitals (*Nationale Raad voor Ziekenhuisvoorzieningen* (2015)) wonders whether the context of a penal institution offers sufficient guarantees for psychiatric treatment, when convicts with psychiatric problems can no longer be interned. Also, it is estimated that it is likely that a convicted prisoner with serious mental problems will be admitted against his will, even if this possibility is no longer explicitly mentioned in the internment act.

Mandatory treatment of psychiatric patients

Historically the internment act and the act on compulsory admission are intertwined. Before the first internment act (1930), mentally ill offenders were not punished, but acquitted and referred to a psychiatric institution, governed at that time by the act on collocation of 1850. This approach was heavily criticized both by judiciary and public opinion because the collocation was often terminated prematurely for pragmatic reasons, without public scrutiny (Smets, Verelst, & Vandenberghe, 2009). Likewise, the current act on compulsory admission allows the physician to stop the psychiatric observation at any time, even prior to the decision of the civil judge (*vrederechter* similar to the *kantonrechter* in the Netherlands) (Rotthier & Van Peteghem, 2010). Besides the fact that the old collocation act was assessed as inadequate to protect society, also medical professionals criticized this procedure. They stressed that their institutions would become unsafe and noted a risk of moral contamination of other patients and the necessity to provide for additional security and confinement measures which hindered the smooth functioning of the institutions (Goethals, 1997). The internment act of 1930 provided for the admission of internees in separate forensic units. However, after the introduction of the act, 20% of the internees (with a rather limited risk profile) were still integrated in the regular mental health care system and the call for specialized forensic care became even louder (Cosyns, 2005). In 2001 the first medium security units (MSUs) were established aiming to provide treatment for internees not needing care in a highly secured hospital, but who are considered too dangerous or unsuitable for a general psychiatric ward or outpatient care. The most dangerous group had to wait till the end of 2014 when the first high security center was established.

Part of the population of internees are convicted internees, which can be considered a special group with a specific need for forensic care. It is important to substantiate statements on this group empirically. However, research into the profile of convicted internees is scarce. The only study on convicted internees was executed in the Walloon institution *Les Marronniers*, and found that the convicted internees compared to regular internees did not differ regarding the psychiatric profile but did differ in terms of risk profile (higher degree of psychopathy and higher scores on risk assessment schemes) (Pham, 2013; Vicenzutto & Pham, 2015). The present study is a follow-up study in a Flemish population.

Present study

In the present study, the profile of a group of convicted internees was compared to that of regular internees, both admitted to a medium security unit in Flanders. In anticipation of the upcoming law reform which enables the transfer of mentally disordered offenders to the regular mental health care system, it was examined whether their profile fits regular care. Based on previous research in Wallonia it was first hypothesized that the psychiatric pathology would be similar in convicted internees compared to regular internees. Second, a higher risk for antisocial behavior was expected based on the judicial profile and scores on risk assessment instruments.

METHOD

Procedure

The population of internees (*N* = 531), treated in one of the three MSUs (Public Psychiatric Care Center in Rekem, the psychiatric center Sint-Jan-Baptist in Zelzate, and the university psychiatric center Sint-Kamillus in Bierbeek) during the period 2001-2010, was divided into two groups: namely those with an administrative internment (convicted internees; *n* = 48) and regular internees (*n* = 483). Demographic, clinical and risk profiles of the two groups were compared.

Regarding clinical variables prior admissions to a general psychiatric unit were analyzed. Psychiatric diagnoses were made cumulatively based on the Diagnostic and Statistical Manual of mental disorders-IV-Text Revision (DSM-IV-TR; American Psychiatric Association, 2000) in consultation with the treating psychiatrists. Intelligence was measured with the *Wechsler Adult Intelligence Scale-III* (WAIS-III; Wechsler, 2000).

The risk profile was assessed through the judicial history and several risk assessment instruments, which have been found to be valid and reliable in previous research (Singh, Grann, & Fazel, 2011). The Psychopathy Checklist-Revised (PCL-R; Hare, 2003) dimensionally (score 0-40) assesses the degree of psychopathic features, with a mean score of 21.5 (*SD* = 6.9) for forensic psychiatric patients in the validation sample (Hare, 2003). Categorically a score of 30 is used as cut-off, while in Europe a score of 25 or above is considered indicative of psychopathy. The Historical, Clinical, Risk management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997) is a structured risk assessment instrument measuring the risk for future violence in terms of low,

medium or high risk. No cut-off score is available when numeric scores (score 0–40) are used. The Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice, & Cormier, 2006) is an actuarial risk assessment instrument for the long term prediction of violence. The total score ranges from –26 to +38, and can be clustered into nine risk categories, with the highest three categories referring to a high risk of recidivism. The mean score in the validation sample was 0.91 (*SD* = 12.9). With the exception of the VRAG (which was scored by the researchers) the data were gathered retrospectively in CPS – and hospital files. Detention periods were registered on the basis of the administrative prison registration system *Detentie Informatie Systeem (SIDIS)*.

The study was approved by the Ethics committee of Antwerp University Hospital.

Participants

The characteristics of the total medium security population (*N* = 531) are presented in the first column of Table 1. The study sample was mainly male (94.9%) and had the Belgian nationality (90.1%). Mean age at first admission to a medium security unit was 36 years (*SD* = 10.82, range = 18.8–73.4). The majority (81.6%) had been treated in a regular psychiatric unit prior to medium security admission. The average number of convictions was 6.3 (*SD* = 5.73, range = 1–40), including 2.4 (*SD* = 1.91, range = 0–11) which were convicted of a violent offense.

The most common DSM-diagnoses were personality disorders (70.6%), substance use disorders (56.7%) and psychotic disorders (43.9%).

The risk of violent recidivism was assessed as high in about one third of the population, based on the PCL-R (33.5%), the HCR-20 (37%) and the VRAG (30.4%).

Data analysis

SPSS version 22 was used for the statistical analyses. Chi-square tests and Fisher Exact tests were used to compare the group of convicted internees to the group of regular internees in case of categorical variables. In case of continuous variables t-tests or Mann-Whitney U tests were used. Data were missing for a substantial portion of the variables; however, excluding all patients with missing values would reduce the sample size to *n* = 133. Therefore, it was decided to conduct the analyses with all available data, resulting in variable sample sizes (Table 1).

RESULTS

Descriptive variables convicted internees

A ministerial decision to internment was made in more than two third (73.3%) of the cases after a conviction for a violent offense, in 24.4% for a property offense and in 2.2% for a traffic offense. The characteristics of the population of convicted internees (*n* = 48) are presented in the third column of Table 1. Convicted internees were exclusively male and had in majority the Belgian nationality (87.5%). They were on average 35 years old (*SD* = 9.31, range = 21.7–67.4) at first MSU admission. The majority (81.6%) had been treated in a regular psychiatric unit prior to medium security admission. The average amount of convictions was 9.5 (*SD* = 7.70, range = 1–38), including 3.3 (*SD* = 2.21, range = 1–11) which were convicted of a violent offense. The most common DSM-diagnoses were personality disorders (72.9%), psychotic disorders (62.5%) and substance use disorders (47.9%). The risk of violent recidivism was assessed as high in about half of the population, based on the PCL-R (55.6%), the HCR-20 (45.2%) as well as the VRAG (58.8%).

Comparison between the two groups of internees

In Table 1 the group of convicted internees was compared to the group of regular internees. Demographically, there were more convicted internees with a history of unstable work ($\chi^2(1) = 5.37$, $p = .02$, *Phi* = .11) in comparison to regular internees.

Clinically, more convicted internees were diagnosed with a psychotic disorder ($\chi^2(1) = 7.43$, $p = .01$, *Phi* = .15). More specifically no differences were found regarding schizophrenia or schizoaffective disorders, but other psychotic disorders such as psychotic disorders not otherwise specified or delusional disorders were found more frequently in convicted internees ($\chi^2(1) = 5.05$, $p = .03$, *Phi* = .15). This difference remained after matching for judicial variables (nature of the offense and amount of convictions) ($\chi^2(1) = 4.36$, $p = .04$).

Table 1. Comparison of regular internees and convicted internees

		Total population		Regular internees (*n* = 483)		Convicted internees (*n* = 48)			
	N	%	*M* (*SD*)	%	*M* (*SD*)	%	*M* (*SD*)	*p*	*Phi/r*
Demographics									
Female sex	531	5.1		5.6		0		.16	
Belgian nationality	527	90.1		90.5		87.5		.46	
Stable work history ≥ 3 y.	484	30.6		32.1		14.6		.02*	.11
Stable relaionship history ≥ 1 y.	474	43.0		44.2		29.7		.09	
Special education	496	23.6		23.5		24.4		.90	
Age at first MSU admission	531		36.5 (10.82)		36.6 (10.96)		35.5 (9.31)	.66	
Judicial history									
Age first conviction	531		24.8 (9.13)		25.1 (9.32)		21.3 (5.89)	.00**	.13
Violent antecedent	531	92.8		92.1		100		.04	.09
Number of convictions	531		6.25 (5.73)		5.9 (5.40)		9.5 (7.70)	.00**	.15
Number of violent convictions	531		2.4 (1.91)		2.3 (1.86)		3.25 (2.21)	.00**	.14
Number of detention periods before MSU	530		4.2 (4.36)		4.1 (4.34)		5.9 (4.2)	.00***	.17
Duration detention periods	530		1691.1 (1824.18)		1565.5 (1708.95)		2951.7 (2404.17)	.00***	.24

Table 1. Comparison of regular internees and convicted internees (continued)

		Total population		Regular internees (*n* = 483)		Convicted internees (*n* = 48)			
	N	%	*M* (*SD*)	%	*M* (*SD*)	%	*M* (*SD*)	*p*	*Phi/r*
Clinical variables									
Prior admission regular psychiatry	521	81.6		81.6		81.6		.97	
WAIS-III score	284		80.5 (16.81)		80.6 (16.80)		80.0 (17.16)	.94	
Psychiatric diagnosis	531								
Psychotic disorders		43.9		42		62.5		.01*	.12
Schizophrenia/schizoaffective		32.8		31.9		41.7		.17	
Other psychotic disorder		11.1		10.1		20.8		.03*	.10
Anxiety and mood disorders		6.6		6.8		4.2		.48	
Personality disorder		70.6		70.4		72.9		.71	
Cluster A		6.8		6.4		10.4		.36	
Cluster B		45.6		45.8		43.8		.79	
Cluster C		3.2		3.3		2.1		1.00	
Nor otherwise specified		15.6		15.3		18.8		.53	
Substance misuse		56.7		57.6		47.9		.20	
Intellectual disability		12.8		12.6		14.6		.70	
Comorbidity both axes		71.2		70.6		77.1		.34	

Table 1. Comparison of regular internees and convicted internees (continued)

		Total population		Regular internees (*n* = 483)		Convicted internees (*n* = 48)			
	N	%	*M* (*SD*)	%	*M* (*SD*)	%	*M* (*SD*)	*p*	*Phi/r*
Risk assessment									
PCL-R									
Total	224		21.1 (6.58)		20.6 (6.47)		24.6 (5.38)	.00**	.20
Factor 1	215		8.6 (3.40)		8.5 (3.43)		9.0 (3.24)	.40	
Factor 2	213		10.3 (3.86)		10.0 (3.87)		12.9 (2.68)	.00***	.25
Facet 1	172		3.1 (2.13)		3.1 (2.17)		3.2 (1.80)	.68	
Facet 2	173		5.6 (1.90)		5.6 (1.95)		6.0 (1.37)	.58	
Facet 3	170		6.1 (2.53)		5.9 (2.55)		7.3 (2.02)	.04*	.16
Facet 4	163		5.5 (2.56)		5.2 (2.53)		7.8 (1.52)	.00***	.31
PCL-R ≥ 30	224	8.9		7.1		22.2		.02*	.17
PCL-R ≥ 25	224	33.5		30.5		55.6		.01*	.17
HCR-20		3.2		3.3		2.1		1.00	
Total	298		24.5 (5.27)		24.3 (5.36)		26.1(3.96)	.09	
H-scale	298		14.1 (3.22)		13.9 (3.27)		15.7 (2.04)	.01*	-.15
C-scale	298		4.7 (1.83)		4.7 (1.82)		5.0 (1.84)	.52	
R-scale	298		5.8 (2.12)		5.8 (2.12)		5.5 (2.23)	.56	
SPJ High	263	38.4		37.5		47.8		.33	
VRAG									
Total	230		8.0 (10.94)		7.3 (10.96)		16.1 (6.74)	.00***	.22
High risk	230	30.4		28.2		58.8		.01**	.17

Note. WAIS-III = Wechsler Adult Intelligence Scale-III; PCL-R = Psychopathy Checklist-Revised; HCR-20 = Historical, Clinical, Risk management–20; SPJ = structured professional judgment; VRAG = Violence Risk Appraisal Guide.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Judicially, convicted internees were younger than regular internees at their first conviction (U = 8540.50, z = -3.01, p = .00, r = -.13). In addition, more convicted internees with prior convictions for a violent offense were found (Fisher, p = .04, *Phi* = .09). Also, convicted internees had on average more convictions for a general offense (U = 8083.50, z = -3.48, p = .00, r = -.15) and for a violent offenses (U = 8446.00, z = -3.20, p = .00, r = -.14) compared to regular internees. Convicted internees were more often incarcerated (U = 7565.50, z = -4.02, p = .00, r = -.17) and remained incarcerated for longer periods (U = 6048.50, z = -5.46, p = .00, r = -.24) prior to their MSU admission. Scores on risk assessment instruments differed significantly between the two groups. Convicted internees had higher PCL-R total scores (t(222) = 3.00, p = .00, r = .20). They scored higher on Factor 2 and Facet 3 and 4, referring to an antisocial life style. Significantly more convicted internees were diagnosed as psychopaths either when a cut-off score of 30 was used (Fisher, p = .02, *Phi* = .17) as when a cut-off score of 25 was used (χ^2(1) = 6.72, p = .01, *Phi* = .17). Further, they had a higher risk profile based on the historical scale of the HCR-20 (U = 2460.00, z = -2.57, p = .01, r = -.15) and the total VRAG-score (t(228) = -3.27, p = .00, r = .22).

DISCUSSION

The procedure for ministerial internment without an adversarial debate, allowing the transfer of mentally disordered convicted prisoners during detention to internment with possibly longer confinement is criticized for the weak legal position of the convicted internee. According to some authors, this procedure violates basic human rights, a view that was contradicted by the Belgian jurisprudence. Belgium seems moreover to be one of the few countries that provide such a procedure. In the new internment law (2014), which is currently under construction (final text is not yet known in April, 2016), the controversial ministerial internment was deleted. The present study examined the implications of this new proposal for the regular mental health care system and thus touches on a broader debate which deals with the connection of regular to forensic care (cf. the current debate in the Netherlands regarding the implementation of the Forensic Care Act and the Compulsory Mental Health Care Act). It was examined whether clinical - and risk profile of the population of convicted internees is compatible with non-categorical care. Medium security convicted internees were compared with regular internees on demographic, clinical and

risk variables. In contradiction to research by Vicenzutto and Pham (2015), who found no psychopathological differences, more convicted internees with a psychotic disorder were found in the present study. Whether these problems were the result of for instance prisonization (*prison psychosis*) could not be traced. In terms of other psychopathology such as schizophrenia, substance related disorders or personality disorders no differences were found, which was in line with the Walloon study (Vicenzutto & Pham, 2015).

Also in line with the study by Pham (2013) a higher risk for recidivism was found in the Flemish convicted internees. This became evident when analyzing the judicial histories (younger age at first conviction, more convictions, more and longer detention periods) and the risk profiles based on risk assessment instruments. For example, the mean VRAG score was high in comparison with a norm group of forensic psychiatric patients (Quinsey et al., 2006). Also one of the major risk factors for (violent) recidivism – a high level of psychopathy – was evident in convicted internees. The PCL-R total score was high in comparison with the norm group of forensic psychiatric patients (Hare, 2003). There was a particular increase in Factor 2 and Facet 4, the behavioral traits of psychopathy that showed the clearest link with an increased risk of recidivism (Leistico, Salekin, DeCoster, & Rogers, 2008; Walters, 2003). The higher scores on the aforementioned characteristics are all the more striking because the comparison group was comprised of regular internees with a medium risk and security level. Although significantly different, the effect sizes remained small. The least we can conclude is that we are dealing with an equally problematic group of patients, comparable to other internees.

Further research is needed to determine whether the present results generalize to the entire group of convicted internees, for whom, as yet, no prevalence rates or profiles are available. Data from the Ministry of Justice show that at present very limited use is made of the measure, with only 19 requests for ministerial internment between September, 2012 and May, 2015. Furthermore, the legislative reform stipulates that psychiatric care to inmates outside the prison system can be provided at the earliest within the modalities of the normal sentence. On the contrary, in the Netherlands, it is possible to transfer a convicted mentally disordered offender temporarily to a forensic psychiatric institution for treatment (Art. 13 placement). Also, other alternatives are considered. As such the proposal for a new law on forensic care stipulates that a criminal judge can impose mandatory treatment, regardless of the location.

Minimally it can be concluded from both studies that the group of convicted internees is one with complex problems. The convicted internee is typically a middle-aged man with a limited social network, severe psychiatric symptoms in combination with addiction and/or personality problems and a lot of criminal convictions. The associated risk profile requires appropriate care with a matching - infrastructural and relational – security level with focus on risk assessment and risk management. The question remains how regular mental health care will handle this new demand for care in an era of de-institutionalization and cost cutting. Another important question is whether the society will be adequately protected when applying a civil placement? This could lead again rather to *risk avoidance* than *risk management*. Other countries, like the United Kingdom therefore provide for extra criminal scrutiny. A placement in a psychiatric hospital during (Section 47/49 Mental Health Act (MHA)) or after detention (Section 37/41 MHA) is coupled with an indefinite restriction order whereby the Minister of Justice decides upon (conditional) release (Rethink Mental Illness, 2015). Important to mention here is the fact that treatment is provided by specialist forensic centers or secure hospitals, with security levels higher than those of regular psychiatric wards.

At a time when patient flows are increasingly intertwined (most of the medium security internees were previously treated in a regular psychiatric unit), it will be important to get the right patient at the right place, both in the interest of the patient as of society at large: “Regular where possible, forensic where necessary”.

CONCLUSION

The abolition of the ministerial internment in the first version of the new internment Act (2014) raises a number of questions. The present study shows that convicted internees can be considered as a complex group, both in a psychiatric and a judicial sense. The abolishment of the measure resolves a number of legal issues related to the legal uncertainty, but creates new – especially clinical – problems. More prisoners will face double statutes during detention which may impede access to psychiatric care. Moreover, the current regular psychiatric care facilities are not equipped to answer to complex forensic care needs. As a result, forensic patients can not only be denied appropriate care but also society risks to become less safe. Alternatives, such as

the opening of (additional) forensic wards for patients with civil statutes and the provision of extra judicial validation when ending the treatment, can be considered.

FOOTNOTES

[1] All offenses for which the Criminal Law sets a minimum penalty of at least 8 days are included.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Berge, Y. Van Den (2008). Het gevreesde artikel 21 van de Wet Bescherming Maatschappij doorstaat de rechterlijke toets, maar hoort het nog thuis in het Belgisch penitentiair recht. *Tijdschrift voor Strafrecht, 5*, 396-397.

Berge, Y. Van Den (2011). Artikel 21 van de Wet Bescherming Maatschappij doorstaat nu ook de toets van het Grondwettelijk Hof. *Tijdschrift voor Strafrecht, 4*, 261-264.

Clerck, A. De, & Steenbrugge, W. Van (2007). Het wetsontwerp internering: het lijden eindelijk verlicht? *Fatik, 114*, 16-20.

Cosyns, P. (2005). *Werkgroep forensisch psychiatrisch zorgcircuit: Syntheseverslag mei 2005*. Brussel: Federale Overheidsdienst Justitie.

Dijkmans, I. (1980). De internering van veroordeelden krachtens een beslissing van de Minister van Justitie: Een maatregel van onbepaalde duur? *Panopticon, 3*, 224-235.

Goethals, J. (1997). De wet tot bescherming van de maatschappij in historisch perspectief. In J. Casselman, P. Cosyns, J. Goethals, M. Vandenbroucke, D. De Doncker, & C. Dillen (Eds.), *Internering* (pp. 11-38). Leuven: Garant.

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist*. Toronto, ON: Multi-Health Systems.

Heimans, H., Beken, T. Vander, & Schipaanboord, A. E. (2015). Eindelijk een echte nieuwe en goede wet op de internering? Deel 1: De gerechtelijke fase. *Rechtskundig Weekblad, 27*, 1043-1064.

Leistico, A. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and Human Behavior, 32*(1), 28–45. doi: 10.1007/s10979-007-9096-6

Moens, I., & Pauwelyn, L. (2012). *Geen opsluiting, maar sleutels tot re-integratie. Voorstellen voor een gecoördineerd zorgtraject voor geïnterneerden*. Brussel: Zorgnet Vlaanderen.

Nationale Raad voor Ziekenhuisvoorzieningen. (2015). *Advies inzake opmerkingen op de wet van 26 juni 1990 betreffende de bescherming van de persoon van de geesteszieke*. Retrieved from http://www.health.belgium.be/eportal/Healthcare/Consultativebodies/Nationalcouncilforhospitalfaci/adviesprogramming2015/19101790_NL#.VkRV9lLluUk

Pham, T. H. (2013, September). *Aspects d'expertise.* Paper presented at the Nationaal congres Bescherming van de Maatschappij, Brussels, Belgium.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Rethink Mental Illness. (2015). *Section 47 49 - Transfer of a sentenced prisoner to hospital*. 2015, from http://www.rethink.org/living-with-mental-illness/police-courts-prison/section-47-49-transfer-of-a-sentenced-prisoner-to-hospital

Rotthier, K., & Peteghem, P. Van (2010). De nazorg als behandelingsmodaliteit in de WPG. In G. Benoit, J. De Fruyt, H. Nys, G. Rommel, G. Steegen, P. Van Peteghem, & J. Van Speybroeck (Eds.), *De bescherming van de persoon van de geesteszieke. Ethsische, medische en juridische perspectieven* (pp. 59-72). Brugge: die Keure.

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499-513. doi: 10.1016/j.cpr.2010.11.009

Smets, H., Verelst, R., & Vandenberghe, J. (2009). Geestesziek en gevaarlijk: Gedwongen opname of internering? Het Belgissche wettelijke kader. *Tijdschrift voor Psychiatrie, 51*(4), 217-225.

Vandevelde, S., Soyez, V., Beken, T. Vander, Smet, S. De, Boers, A., & Broekaert, E. (2011). Mentally ill offenders in prison: The Belgian case. *International Journal of Law and Psychiatry, 34*(1), 71-78. doi: 10.1016/j.ijlp.2010.11.011

Vicenzutto, A., & Pham, T. H. (2015). Les condamnés qu'on interne: plus fous? plus malades. *Annales Médico-Psychologiques, 173*, 629-630.

Walters, G. D. (2003). Predicting institutional adjustment and recidivism with the Psychopathy Checklist factor scores: A meta-analysis. *Law and Human Behavior, 27*(5), 541–558. doi: 10.1023/A:1025490207678

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence* (Version 2). Burnaby, BC: Simon Fraser University.

Wechsler, D. (2000). *WAIS-III Nederlandstalige bewerking. Afname en scoringshandleiding*. Lisse: Swets & Zeitlinger.

Weis, K. (2010). Gewikt en gewogen: De internering opgelegd door de Minister. *Rechtskundig Weekblad, 73*(27), 1114-1124.

Critical incidents and judicial response during medium security treatment

Accepted pending minor revisions

5

Jeandarme, I., Wittouck, C., Vander Laenen, F., Pouls, C., Heimans, H., Oei, T. I., & Bogaerts, S.

ABSTRACT

This study examined inpatient incidents in Flemish forensic medium security units and analyzed the subsequent judicial reactions to these incidents. During medium security treatment, incidents were reported for more than half of the participants. The most frequently registered incidents were non-violent in nature, such as absconding and treatment non-compliance. The base rate for physically violent incidents was low. Although crime-related incidents during medium security treatment were rarely prosecuted and adjudicated, the base rate of revocation – and hence drop-out from treatment – as a result of these incidents was high.

INTRODUCTION

Forensic psychiatric patients have traditionally been stigmatized as more violent, more difficult to treat and less compliant than other patients (Lamb, Weinberger, & Gross, 1999; Schanda, Stompe, & Ortwein-Swoboda, 2009). As a result, general psychiatric institutions are reluctant to treat these patients (Muller-Isberner, 1996) and local communities are opposed to the presence of forensic units because they are concerned for the public safety, for example after absconding (Gradillas, Williams, Walsh, & Fahy, 2007).

Critical incidents during (medium security) forensic psychiatric treatment

A consensus definition of what is understood as a critical inpatient incident is non-existing (Gradillas et al., 2007). Some scholars refer to all inpatient incidents as serious rule violations. For example, Main and Gudjonsson (2006) found that 57% of forensic inpatients reported at least one serious rule violation, defined as absconding, using illicit drugs or consuming alcohol, being violent to staff or patients, damaging property, and fire setting. In the present study, violent and non-violent incidents were separately addressed. In addition, in order to examine judicial responses, it seemed relevant to separate critical incidents falling under offense coding categories from critical incidents referring to treatment interfering behavior.

In the literature, inpatient violence or violence occurring during forensic psychiatric treatment has received substantial attention, in first instance because it affects the stability of an institution, the staff turnover and also because it has a negative impact on the therapeutic process (Gow, Choo, Darjee, Gould, & Steele, 2010; Quanbeck, 2006). However, to determine the number and characteristics of violent incidents in these forensic psychiatric settings, the definition of a violent (or aggressive) incident should be carefully scrutinized since – besides physical violence towards others – also verbal violence and/or violence towards self or objects can be included (Alia-Klein, O'Rourke, Goldstein, & Malaspina, 2007; Cullen et al., 2015; Daffern, Duggan, Huband, & Thomas, 2008; Decaire, Bedard, Riendeau, & Forrest, 2006; Gow et al., 2010; Gudjonsson, Rabe-Hesketh, & Wilson, 2000). Unfortunately, the proportion of each subtype of violence is not always separately described in research (Daffern et al., 2008; Decaire et al., 2006). In some studies, a distinction was made between physical assaults and other types of violence

(Gudjonsson et al., 2000). Gudjonsson et al. (2000), who rated the severity of all incidents at a medium secure unit during a time interval of 16 years, found that 47% of the incidents comprised physical violence against persons (76% of these incidents caused however no injury or pain) and 53% comprised verbal violence, damage to property, arson and self-injury. In the study of Gow et al. (2010), only 17.2% of the incidents on a medium secure unit comprised physical violence whereas, 21.8% comprised verbal abuse and 17.6% threats. Overall, a literature review found that the patient base rate for general aggression in forensic settings was 47.7% and the base rate for physical violence was 19.1% (Bowers et al., 2011), with higher percentages for physical violence in single studies, e.g., 28.2% in Gow et al. (2010). Sexual violence and arson were less frequently observed (Blattner & Dolan, 2009; Gow et al., 2010).

Next to physical and verbal violence, also non-violent incidents are reported in forensic units treating offenders with a mental illness, such as violations of hospital rules, treatment non-adherence, the use of alcohol and illicit drugs, and absconding (Abidin et al., 2013; Blattner & Dolan, 2009; Gow et al., 2010; Hillbrand, 1995). Patient base rates of absconding ranged from 4.8% to 21.7% (Blattner & Dolan, 2009; Cullen et al., 2015; Gow et al., 2010). Absconding was only in a minority of cases accompanied by offenses (Gradillas et al., 2007). A recent UK prospective cohort study of medium and low secure forensic psychiatric wards showed that recent verbal aggression and recent substance use was predictive of absconding (i.e., failure to return from leave, incidents of escape, and absconding whilst on escorted leave) (Cullen et al., 2015). Research focusing specifically on non-violent incidents during forensic psychiatric treatment is scarce (Blattner & Dolan, 2009; Gradillas et al., 2007). Yet, in the study of Decaire et al. (2006) on a minimum security forensic unit in a medium secure psychiatric hospital in Canada, 42.3% of the recorded incidents on the unit concerned non-violent incidents such as absconding and violation of ward rules. These authors also hypothesized that staff is more discrete in reporting non-violent incidents than violent incidents, so this number could be an underestimation.

Studying non-violent and (verbally) violent incidents is nevertheless important because both can affect the treatment process by resulting in early treatment termination. Drop-out from treatment can have serious consequences since research has shown that drop-out is associated with recidivism (McMurran & Theodosi, 2007; Olver & Wong, 2009). In a medium secure unit for personality disordered male offenders in the UK, 37% of the patients was expelled from

treatment due to rule-breaking behavior, verbal assault, physical assault or drug offenses and 35.8% was transferred back to prison because they were not actively engaging with the treatment program (McCarthy & Duggan, 2010). Thus, besides rule-breaking behavior, treatment non-engagement can be another reason for drop-out. Another study found that women with low treatment engagement were involved in more adverse incidents, including both physical and verbal aggression during their medium security stay (Abidin et al., 2013; Blattner & Dolan, 2009; Gow et al., 2010; Hillbrand, 1995). Treatment non-engagement can thus be regarded as a specific form of a non-violent incident that is related to poor treatment outcome.

Reporting incidents to the legal authorities

To the best of our knowledge, empirical studies specifically focusing on the reporting of violent and/or non-violent (e.g., theft or drug-related offenses) crime-related incidents during forensic psychiatric treatment to police or judicial authorities are quasi non-existent. Only one study briefly described the characteristics of 41 successful prosecutions of 30 Rampton Hospital inpatients for violent offenses against staff and indicated a need for more research in this area (Clark, McInerney, & Brown, 2012). Van Leeuwen and Harte (2011) described judicial reactions towards violent patients in general psychiatric services. In the Netherlands, only one out of four incidents of institutional physical violence towards mental health professionals was reported to the police (Harte, Van Leeuwen, & Theuws, 2013).

Jurisprudential, therapeutic, as well as ethical issues seem to form barriers to report incidents or to prosecute patients, particularly in case where patients are already detained in forensic settings. Ambivalence exists within mental health services to report violent incidents to the criminal justice system. Many mental health professionals are unwilling to consider a course of action that may be punitive for the patient involved, reasoning that such an action might hamper the treatment relationship (e.g., by breaching patients' confidentiality) and it might harm the patient. These consequences seem inconsistent with their role as caregiver. Furthermore, even when incidents are reported to police services, judicial authorities appear to be reluctant to prosecute and convict those patients (Dinwiddie & Briska, 2004). Several barriers can hamper prosecution such as an inability to collect the necessary information concerning the crime and the intention of the perpetrator (due to confidentiality issues) as well as doubts about whether

the threshold for prosecution and eventual outcome at court is justified (Clark et al., 2012). For instance, in the aforementioned Dutch study, only 10% of the physical violent incidents reported to the police were brought to court (Harte et al., 2013). The study by Clark et al. (2012) in a forensic hospital showed that there was no clear presumption that patients would be prosecuted for assaults on staff, despite the zero tolerance policy for this type of assaults.

Present study

The present study adds to the scant literature on critical incidents in forensic psychiatric settings by examining incidents occurring during treatment in a sample of medium security forensic patients. First, the period prevalence and characteristics of violent as well as non-violent incidents and the patients involved were examined. Second, the judicial reaction to the reported (crime-related) incidents was investigated. In line with previous research findings it was hypothesized that 1) physically violent incidents would be less prevalent than verbally violent and non-violent incidents and 2) crime-related incident reports would be rarely prosecuted and adjudicated.

MATERIAL AND METHODS

Setting and participants

This multicenter study was conducted at three medium security units located in the Flemish communities of Bierbeek, Zelzate and Rekem. Medium security units provide a treatment setting for patients found not guilty for reason of insanity (NGRI or *interned*) who do not require care in a high secure hospital, but who are considered unsuitable for a general psychiatric ward or outpatient care (see Jeandarme, Habets, Oei, and Bogaerts (2016) for a description of the Belgian forensic psychiatric system). Referral to a medium security unit is provided under conditional release and is linked to specified conditions. The imposed conditions can be divided into orders (following hospital rules, being compliant with the treatment and directions of the probation officer) and prohibitions which typically include both general conditions (not committing new offenses, not using drugs) and individualized conditions (e.g., restraining order). The supervision of the abovementioned conditions is done by a regional court, the multidisciplinary Commission of the Protection of Society (CPS), chaired by a judge. A public prosecutor is present at the

hearings and advices the CPS on (conditional) release, but only the official members of the CPS (a judge, a psychiatrist and a lawyer) take part in the formal decision making process. Violent (and other) incidents occurring during the treatment are reported to the CPS on a regular basis by the medium security units and/or the probation officer. When conditions are breached (either due to new offenses or due to other incidents such as absconding), the public prosecutor can decide to re-incarceration in prison, which implies a revocation of the conditional release. The reason for this re-incarceration is that, at the time of the study, high security forensic units were non-existent in Flanders, thus medium security patients could not be transferred to a high security unit and transfers to another medium security occurred only in rare occasions (Jeandarme et al., 2016; Vandevelde et al., 2011).

Most of the study participants were male patients (94.9%, *n* = 504); only a minority were female patients (5.1%, *n* = 27). The majority of the study sample had the Belgian nationality[1] (90.1%, *n* = 475). Participants were found NGRI for a variety of offenses: violent offenses (77.2%, *n* = 410), property offenses (18.6%, *n* = 99), drug-related offenses (2.3%, *n* = 12), sexual non-violent offenses (0.8%, n = 4) and other offenses (1.1%, *n* = 6).[2] Prior contact with the criminal justice system was common; only a minority (15.6%, n = 83) was registered as first offender. The average amount of convictions for a general offense was 6.3 (*SD* = 5.73, range = 1–40), among which 2.6 (*SD* = 1.87, range = 1–11) for a violent offense. Preceding the admission to the medium security unit, the population was incarcerated for on average 4.6 years (*SD* = 4.99, range = 0–39 years).

Mean age at first admission to a medium security unit was 36.5 years (*SD* = 10.82, range = 19–73 years). More than half of the study population (52.9%, *n* = 281) was admitted to a medium secure unit more than once. The average length of these admissions[3] was 676.4 days (*SD* = 505.90; range = 8–2729 days). Five patients died during treatment and about 10% of the internees (*n* = 65) were still in treatment at the end of the study. Of the remaining group (*n* = 461), two-thirds (68.3% *n* = 315) completed treatment, while a third (31.7%, *n* = 146) dropped out prematurely.

The most common cumulative diagnoses according to the Diagnostic and Statistical Manual of mental disorders-IV-Text Revision (DSM-IV-TR; American Psychiatric Association, 2000) were personality disorders (70.6%, *n* = 375), substance use disorders (56.7%, *n* = 301), psychotic disorders (43.9%, *n* = 233), and intellectual disabilities (23.0%, *n* = 122).

Outcome measures

Critical incidents

Critical incidents[4] were categorized into non-violent and violent incidents (see Table 1). Non-violent incidents included: 1) treatment non-compliance referring to not adhering to treatment rules (e.g., not engaging in treatment plans, refusing prescribed medication and drinking alcohol during treatment), 2) breach of judicial conditions referring to not adhering to conditional release conditions imposed by the CPS (e.g., ignoring a contact order or not attending appointments with the probation officer), 3) absconding referring to absconding from (un)supervised leave or absconding without permission to leave the premises, 4) drug-related incidents referring to usage or dealing of illegal drugs, 5) property incidents (further subdivided into arson in an uninhabited property and other property incidents such as thefts), 6) sexual non-violent i.e., hands-off offenses, and 7) other incidents further subdivided into traffic offenses and other incidents such as damage to property.

Violent incidents referred to the intentional use of physical force or power – threatened, attempted or actual – against another person. Verbal violence referred to threatened violence; physical violence to attempted or actual interpersonal violence. Physical violence was further subdivided into attempted manslaughter, sexual violence, arson in a habited property, property offenses with violence and other forms of violence (e.g., assault).

Judicial reaction to the incidents

When an incident report in a CPS file consisted of multiple critical incidents, the report was classified according to the nature of the most serious incident. CPS decisions following incident reports were divided into 1) no judicial reaction, 2) modification of judicial conditions, and 3) revocation of the conditional release. In order to determine which incidents were prosecuted and trialed, crime-related incidents were separated from the reported incidents. Crime-related incidents were defined as incidents coded under offending categories of the Belgian penal code, whether or not they led to further prosecution or sentencing, and consisted of violent, sexual non-violent, property, drug related, and other offenses.

Table 1. Nature of incidents (event- and person-related)

	Incidents		Patients involved in incidents	
	N	%	*n*	%
Incidents total	955	100	303	57.1
Non-violent incidents	766	80.2	274	51.6
Absconding	308	32.3	158	29.8
Non-compliance	245	25.7	158	29.8
Drug incidents	123	12.9	69	13.0
Property offenses	35	3.7	29	5.5
Property offense without violence	33	3.5	28	5.3
Arson (goods, uninhabited property)	2	0.2	1	0.2
Other offenses	29	3.0	26	4.9
Offense not otherwise specified	24	2.5	23	4.3
Traffic offense	5	0.5	4	0.8
Breach of judicial conditions	22	2.3	20	3.8
Sexual hands-off	4	0.4	4	0.8
Violent incidents	189	19.8	115	21.7
Verbal violence	93	9.7	73	13.7
Physical violence	96	10.1	69	13.0
Other violence	72	7.5	54	10.2
Arson occupied property	9	0.9	9	1.7
Property offense with violence	8	0.8	6	1.1
Sexual hands-on offense	5	0.5	5	0.9
Attempted manslaughter	2	0.2	2	0.4

Procedure

Data were collected for all patients treated in a medium security unit in Flanders in the period 2001–2010. All patients were asked for passive consent. Eleven patients refused, leading to a total sample of 531 patients (Zelzate: 41.1%, *n* = 218; Bierbeek: 28.8%, *n* = 153, and Rekem: 30.1%, *n* = 160). Data on person-related information of patients involved in incidents was sought from hospital and judicial CPS files. Event-related information of incidents was gathered by reviewing incident reports in the Flemish judicial files from the four CPSs (CPS Antwerp: 45.4%, *n* = 241; CPS Ghent: 38.2%, *n* = 203; CPS Leuven: 13.2%, *n* = 70; CPS Vorst: 2.8%, *n* = 15, and CPS unknown: 0.4%, *n* = 2). The judicial reaction to the incident reports was analyzed by reviewing CPS decisions and conviction data derived from the Central Criminal Records of the Justice Department. In case

of revocation of the conditional release and thus re-incarceration, the duration of the detention period was analyzed from the prison registration system of the Justice Department.

The study was approved by the Ethics committee of Antwerp University Hospital on January 24, 2011.

SPSS Version 22 was used for the descriptive statistical analyses. Chi-square test was used to compare revocation rates between incident reports for violent and non-violent incidents.

RESULTS

Event-related incidents

In total, 957 incidents were recorded in the CPS files. For two incidents (0.2%) there were missing data concerning the nature of the incident. The majority of the incidents were non-violent incidents (80.2%, n = 766) and a minority were violent incidents (19.8%, *n* = 189) (Table 1). The most common non-violent incident was absconding (40.2%, *n* = 308/766). Absconding occurred while on the hospital premises (74.0%, *n* = 228/308) and during unsupervised (23.4%, *n* = 72/308) or supervised leave (2.6%, *n* = 8/308). They lasted on average 24.8 days[5] (*SD* = 177.06, range = 1–2833 days), but were terminated within two days in 69% of the cases (*n* = 207/308). In one out of five abscondings (20.1%, *n* = 62/308) an additional crime-related incident occurred (see further). Furthermore, other non-violent incidents were: non-compliance (32.0%, *n* = 245/766), drug-related offenses (16.1%, *n* = 123/766) such as using (*n* = 122), using and selling (*n* = 3) or selling illegal drugs (*n* = 1), property offenses without violence such as theft, fraud, and arson in an uninhabited property (4.6%, *n* = 35/766), other offenses such as damaging property or traffic offenses (3.8%, *n* = 29/766), breach of judicial condition(s) (2.9%, *n* = 22/766) and hands-off sexual offenses (0.1%, *n* = 4/766).

Almost half of the violent incidents (49.2%, *n* = 93/189) consisted of verbal violence (e.g., threats and stalking), whereas the other half (50.8%, *n* = 96/189) included physical violence (e.g., battery and sexual assault). Sixteen cases of the physical violent incidents can be considered very serious, e.g., two attempted manslaughter, nine arsons and five sexual hands-on offenses.

One out of five crime-related incidents (22.4%, *n* = 85) occurred during absconding. More specifically, these incidents concerned drug-related offenses (*n* = 32), physical violent offenses (*n*

= 18), property offenses (*n* = 17), verbal violent offenses (*n* = 9) and other offenses (*n* = 9). Among the physically violent offenses which occurred after absconding there was one attempted manslaughter and one sexual assault against an adult victim.

Event-related incident reports

In the CPS files, a total of 680 incident reports were found, including the 955 incidents described earlier (see 3.1). Two incident reports (0.1%) with missing data were excluded from the analyses. Thus, the majority of incident reports consisted of non-violent incident reports (77.9%, *n* = 528/678) and a minority of violent incident reports (22.1%, *n* = 150/678). Most incident reports consisted of incidents which occurred on the unit or the hospital premises (79.8%, *n* = 541/678); a minority occurred during unsupervised (18.3%, *n* = 124/678) or supervised leave (1.9%, *n* = 13/678). Table 2 shows the distribution of the different incident reports stratified per setting.

Table 2. Nature of incident reports (event-related) total and stratified per setting

	IR total	IR on the unit/premises	IR during unsupervised leave	IR during supervised leave
Total	678	541	13	124
Non-violent IR	528	412	10	106
Absconding	246	183	7	56
Drug related IR	112	92	0	20
Non-compliance	109	91	1	17
Property offenses IR	25	20	1	4
Breach of CPS conditions	16	9	0	7
Other offenses IR	16	14	2	0
Sexual hands-off IR	4	3	1	0
Violent IR	150	129	3	18
Physically violent IR	84	72	3	9
Verbally violent IR	66	57	9	0

Note. IR = incident report; CPS = Commission of the Protection of Society.

Judicial reaction to event-related incident reports

In 50.4% of the reports (*n* = 342/678) the medium security treatment was continued without imposing additional conditions although typically in these cases the internee would be cautioned by the chairman at a CPS hearing. In 1.3% (*n* = 9/678) the treatment was continued after imposing

additional conditions. In 48.2% of the reports (*n* = 327/678) the prosecutor decided to detention and the conditional release was subsequently revoked by the CPS. The most common reasons for re-imprisonment were 1) violent offenses (33.6%, *n* = 110/327), 2) abscondings (24.5%, *n* = 80/327) and 3) non-compliance (21.1%, *n* = 69/327). Revocation occurred significantly more often when the report was characterized by a violent compared to a non-violent offense ($\chi^2(1) = 48.61$, $p < .001$). Table 3 shows the type of the reports in more detail, qualified according to the most serious incident.

Table 3. Decisions of the Commission of the Protection of Society related to the most severe type of incident reported

	N	Revocation conditional release *n* (%)	Modification judicial conditions *n* (%)	No judicial reaction *n* (%)
Total	678	327 (48.2)	9 (1.3)	342 (50.4)
Non-violent IR	528	217 (41.1)	7 (1.3)	304 (57.6)
Absconding	246	80 (32.5)	0	166 (67.5)
Drug related IR	112	38 (33.9)	4 (3.6)	70 (62.5)
Non-compliance	109	62 (56.9)	2 (1.8)	45 (41.3)
Property offenses IR	25	18 (72)	0	7 (28)
Breach of CPS conditions	16	7 (43.8)	0	9 (56.3)
Other offenses IR	16	10 (62.5)	1 (6.3)	5 (31.3)
Sexual hands-off IR	4	2 (50)	0	2 (50)
Violent IR	150	110 (73.3)	2 (1.3)	38 (25.3)
Physically violent IR	84	66 (78.6)	1 (1.2)	17 (20.2)
Verbally violent IR	66	44 (66.7)	1 (1.5)	21 (31.8)

Note. IR = incident report; CPS = Commission of the Protection of Society.

In almost a quarter of the incident reports (24.2%, *n* = 79), it was not clear who initiated the request to end treatment. From the reports that did register who requested the discontinuation of treatment, it became clear that most requests were made by the medium security unit (74.6%, *n* = 185/248). In one out of ten reports (10.9%, *n* = 27/248), the patient requested to be transferred to prison and in the remaining cases (14.5%, *n* = 36/248) this was the public prosecutor.

As mentioned before, the crime-related incident reports were extracted from the other incident reports in order to determine the judicial response. Almost half of the incident reports were crime-related (45.3%, *n* = 307). In 58% (*n* = 178) of the crime-related incident reports, the conditional release was revoked. Formal court procedures other than CPS rulings[6] were only completed in 15 incident reports[7] (4.9%). These procedures resulted in 18 convictions, consisting of 11 new NGRI verdicts (internments), four convictions to a prison sentence (ranging from 6 to 180 months) and three fines.[8] These sentences were pronounced after incident reports characterized by violent offenses (38.9%, *n* = 7), traffic offenses (33.3%, *n* = 6), and property offenses (27.8%, *n* = 5). Interestingly, all but one of the adjudicated incident reports occurred during leave outside of the unit.

Person-related incidents

The number of patients with any incidents was further divided into violent and non-violent incidents. The number of incidents per patient is shown in Figure 1. As can be seen, the distribution was positively skewed with the majority of the patients involved in no or few incidents and less than 10% of the patients involved in three or more incidents (Figure 1).

Slightly more than half of the population (57.1%, *n* = 303) was involved in at least one incident (*M* = 1.80, *SD* = 2.55, range = 0–18). About 20% of the population (21.7%, *n* = 115) was involved in at least one violent incidents (*M* = 0.35, *SD* = 0.80, range = 0–5) and about half (51.6%, *n* = 274) was involved in at least one non-violent incidents (*M* = 1.44, *SD* = 2.21, range = 0–14). Table 1 shows more in detail that one-third of the population absconded (29.8%, *n* = 158) and one-third did not comply with treatment rules (29.8%, *n* = 158). A minority (3.8%, *n* = 20) breached judicial conditions imposed by the CPS. Furthermore, 13% (*n* = 69) committed drug-related offenses, 5.5% (*n* = 29) property offenses, and 4.9% (*n* = 26) other offenses, mainly damaging property. Four internees committed a sexual hands-off offense. Only a few patients were reconvicted for a general (4.0%, *n* = 12) or violent offense (2.3%, *n* = 7) during medium security treatment.

Two-thirds of the patients involved in an incident (71.6%, *n* = 217) was reincarcerated at least once during medium security treatment due to revocation. On average, the cumulative detention period lasted 674.8 days (*SD* = 568.53, range = 3–2807 days). In 34.1% of the re-

incarcerations, treatment was interrupted temporarily (*n* = 74/217) and in 65.9% of the cases (*n* = 143/217) revocations ultimately caused premature drop-out from medium security treatment.

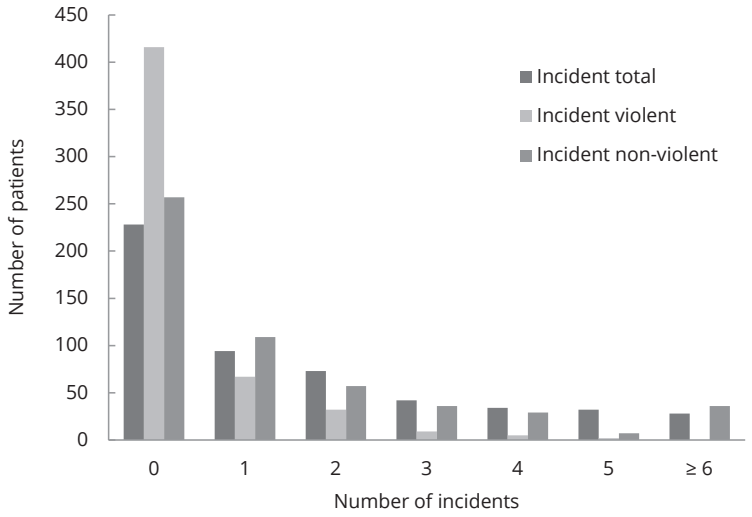


Figure 1. Distribution of number of total, non-violent and violent incidents per person.

DISCUSSION

The aims of the present study were to determine the period prevalence and characteristics of recorded incidents during medium security treatment and the subsequent judicial reaction. Most incidents were non-violent, with a higher than expected number of abscondings. As expected, revocation rates were high and reconvictions rates following crime-related incidents were low.

Non-violent and non-crime-related incidents are most prevalent

Reported incidents consisted mostly of non-violent incidents, such as abscondings, problems with (treatment) compliance and drug-related incidents. In comparison to previous research (Blattner & Dolan, 2009; Gow et al., 2010) the amount of persons that absconded in our study was 3.5 fold higher. One possible explanation could be that patients are granted privileges to leave the unit (un)supervised relatively soon after admission in order to attend treatment

sessions that are held on the hospital premises. Next, (un)supervised leaves are also used to assess compliant behavior of the patients before they are referred to less secure units. In line with previous research findings, however, abscondings were only in 20% of the cases associated with crime-related incidents. Gradillas et al. (2007) found an even lower proportion; with less than one out of ten abscondings resulting in a crime-related incident (7.3%, Gradillas et al., 2007). Non-compliance was the second most registered incident. No studies were found using the same or similar definitions of non-compliance. Studies range from very broad definitions including all inpatient infractions (e.g., Main & Gudjonsson, 2006), to narrow definitions referring to medication compliance (e.g., Alia-Klein et al., 2007). In the present study, non-compliance referred to treatment engagement as well as to adherence to hospital rules but excluded other incidents such as abscondings or violence. In general, there is a thin line between treatment non-compliance and non-engagement, with inconsistent definitions and assessments generating confusion to these concepts (Holdsworth, Bowen, Brown, & Howat, 2014). Drug-related incidents (mostly illicit drug use) were the third most frequently registered incidents, which were found in more than half of the study population. Together with the high percentage of substance misuse diagnoses this finding highlights the importance of focusing on ongoing substance misuse treatment or access to addiction services during medium security treatment.

The minority of the incidents (19.8%) included violent incidents. As hypothesized, and in line with previous research (Gow et al., 2010; Gudjonsson, Rabe-Hesketh, & Wilson, 1999), a considerable amount of the violent incidents concerned verbal violence (49%). The other half of the violent incidents concerned physical violence. In 16 cases (8.5% of the violent incidents or 1.7% of the total number of incidents), the physical violence was very serious[9]: it concerned two attempted manslaughters, nine arsons and five sexual hands-on offenses. Although two incidents were of very serious violence during absconding (one attempted manslaughter and one sexual assault), the total number of violent incidents (incuding verbal violence) that occurred during absconding (n = 27) was small given the fact that we studied three units during an extensive study period of 10 years. Our results were in line with the study of Gradillas et al. (2007), which showed that, despite concern from local residents about the implementation of secure units, the MSUs did not have a significant impact on serious crime rates in their local communities. The person and event base rate for physical violence on the units in the present

study was also low (i.e., 13% and 18.1% respectively) compared to the base rates found in a meta-analysis (i.e., 19.1% and 39.7% respectively) of studies in forensic psychiatric settings (Bowers et al., 2011). Further, the low percentage of patients involved in verbal and physical violence was remarkable, particularly given the fact that most of the patients had a history of (sexual) violent convictions. On the one hand, this can be explained by the fact that apart from individual factors, situational factors such as management approaches and staff characteristics can lead to reduced levels of inpatient violence (Gadon, Johnstone, & Cooke, 2006). On the other hand, it can also be expected that verbal violence was underreported by staff, if these behaviors are considered less serious by staff or if staff is tolerant towards these incidents (Gow et al., 2010).

Incident reports lead to re-imprisonment

In about half of the incident reports (48.2%), the conditional release was revoked and the patients returned to prison. As could be expected, violent offenses were more often followed by imprisonment as compared to non-violent offenses. Interestingly, more than 40% of the revocations were caused by non-crime-related incident reports, such as absconding and non-compliance. Given that the internment measure is an indeterminate measure and there are not so many treatment options, the imprisonment periods following revocation were lengthy, lasting on average 1.8 years. Furthermore, they interrupted or stopped the medium security treatment altogether. As Wormith and Olver (2002) noted, "getting higher risk offenders into treatment is one thing, and keeping them there is another" (p. 449). Non-completion also calls into question whether the level of security in the present study was matched to the population at hand and whether the treatment was responsive to the offenders' needs. Besides hospital-initiated expulsion, it is interesting to note that some of the re-incarcerations (10.9%) were requested by the patients themselves. Internees residing in medium security units may feel too much pressure in treatment and a lack of control compared to when residing in prison (To, 2015).

International research findings show it is not unusual for medium security units to transfer patients back to high security units (Blattner & Dolan, 2009; Davoren et al., 2012; Gow et al., 2010; Hollin et al., 2013). In Flanders however internees cannot be transferred to a high security unit since such a facility was not available at the time of the study. Revocation in Flanders thus accords to re-incarceration, where access to psychiatric care is very limited or even absent. The

European Court of Human Rights has repeatedly convicted Belgium for the imprisonment of internees. The Court ruled that incarcerating persons with an *unsound mind* is only lawful when the facilities used for this purpose are adjusted to their disabilities and needs, and that, if this is not the case, the incarceration can be regarded as torture or an inhuman and degrading treatment or punishment (Heimans, Vander Beken, & Schipaanbord, 2015; Meysman, 2016). In light of the lack of treatment possibilities for mentally ill offenders in Belgian prisons, it would be advisable to limit the number of re-incarcerations to an absolute minimum (e.g., after a very serious physically violent incident). This could be achieved by 1) training of staff in preventing and managing violence, taking into account individual as well as situational risk factors (Decaire et al., 2006; Gadon et al., 2006) and 2) developing an integrated forensic psychiatric care system wherein medium security units refer patients mutually (e.g., in case of drug use) or to a high security unit (e.g., in case of a physically violent incident) depending on the nature of the incident. When an incident occurs on a (forensic) psychiatric unit, a reaction must follow. This reaction should however be adequate and embedded in forensic psychiatric treatment (Quanbeck, 2006). Since in 2014 the first high risk forensic psychiatric center was opened in Flanders, the development of a formal collaboration between the medium security units and the current and future forensic psychiatric centers in Flanders regarding crisis or time-out admissions and high security transfers is of utmost importance. Unfortunately, until now, few initiatives have been taken. Although the medium security units decided in September 2013 on a mutual agreement for time-out placements in case of crisis situation, this is rarely being used. Referrals to the recently implemented high security unit are only possible after incarceration and there is already a long waiting list. Only very exceptionally and for very urgent reasons, the Ministry of Justice and/or the CPS can decide to transfer an inmate immediately to the high security unit.

Avoiding drop-out from treatment is very important for several reasons. First because there is a higher risk for recidivism in case of treatment non-completion. Second, chances for further treatment within another setting may be compromised. Carr et al. (2006) explain the high number of forensic treatment drop-out by a poor treatment preparation during detention. As was shown in the present study, detention periods prior to the first medium security treatment were extensive. Research revealed that inmates adopt attitudes that are adaptive in correctional settings (such as distrust towards staff, behaviors such as intimidating and dissimulation of

symptoms), but become maladaptive once released (Rotter, Carr, Magyar, & Rosenfeld, 2011). Moreover, inmates with a mental disorder tend to have even greater difficulties coping with this adaptation (Carr et al., 2006). In view of the lack of sufficient treatment in Belgian prisons, it is advised to start at least pre-therapeutic counseling in prison to reduce the adaption process (or the culture shock) for mentally ill offenders from a prison to a treatment setting (Carr et al., 2006). This has been partly developed in the last years in Flanders (Stassen, Habets, Mertens, De Laender, & Jeandarme, 2014).

Crime-related incidents only seldom lead to new convictions

Crime-related incidents seldom led to new convictions (4.9%). This finding is in accordance with research in Dutch general psychiatry, which revealed that only 10% of the physical violent incidents reported to the police were brought to court (Harte et al., 2013). Another study in New Zealand suggested that referral of prosecution for violent patients remains rare and arbitrary, despite being increasingly mentioned as an option for staff (Kumar, Fischer, Ng, Clarke, & Robinson, 2006). To the best of our knowledge, there are no studies investigating whether prosecution and conviction rates in forensic settings differ significantly from those in general psychiatric settings. One study in a high security hospital in England found that successful prosecutions do occur in forensic settings, although also at a low rate (e.g., 41 successful prosecutions for assaults on staff or slightly over 10% of total staff assaults) (Clark et al., 2012). A low reconviction rate can be the consequence of a low rate of reporting, prosecuting or sentencing. As has been shown in previous research, the decision to report for example a violent incident to police services can be influenced by different factors, including the influence of the institute's policy and the attitudes of colleagues (Harte et al., 2013). In addition, confusion among treating psychiatrists concerning conflicts of interest and confidentiality issues is likely. Clark et al. (2012) showed that a consistent and consequent approach within a local agreement between psychiatrists and the investigating and prosecuting authorities can overcome some of the barriers. In the present study, incidents were reported – mostly indirectly through probation officers – to the supervising CPS, but it was not known if this resulted in a formal complaint.

Even if reported, police enquiries might be hampered due to limited cooperation both by patients due to their psychiatric symptoms as well as by clinicians due to confidentiality issues.

Further, prosecuting psychiatric inpatients is often seen as poor use of limited resources and it could be argued that protection of the society is already accomplished by hospitalization while the risk to other patients and staff is perceived as being "part of the job" (Dinwiddie & Briska, 2004). The present study did not have access to data on police enquiries or prosecution practices. Information was available on the sentencing level, where it became obvious that few crime-related incidents were prosecuted and subsequently sentenced. The present study revealed that most crime-related incidents were less serious in nature, which might be another reason to renounce prosecution. Furthermore, in the specific case of the Belgian internment measure, the lack of prosecution may be due to the mandatory supervision of the CPS, which allows the prosecutor to re-incarcerate the internee within a rather flexible procedure without contradictory debate that is immediately carried out, thereby avoiding the lengthy and lingering procedure of a new trial, which would in most cases result in yet another internment measure. However, these re-incarcerations can be considered as severe punishments for crime-related incidents which are often not very serious in nature. Patients are sent back to prison, without any formal trial, which further hampers their chances of returning to a forensic unit (and, in a later stage, to general psychiatry). These re-incarcerations might further increase the negative stereotypes held by the general psychiatry and the distrust of forensic patients in (forensic) mental health services. Furthermore, because there is no prosecution and conviction, the rights of the potential victims are not taken into account by the criminal court. Currently, legal changes concerning the internment measure, with more procedural rules and due process, are under way and will be implemented in October, 2016.

Limitations

This study was based on incidents reported to the CPS. Since there is neither formal obligation nor standardized reporting in Flemish forensic institutions to the CPS court because of a lack of a general protocol regarding this matter, the *dark number* of incidents remains unknown. However, it is likely that incidents – especially more serious ones – will be reported anyway since the treatment of internees is being supervised by this court. Furthermore, the attitude of Flemish hospital staff towards violent incidents was not evaluated. It is not known, for example, if Flemish

nurses are tolerant towards (verbally) abusive behavior, resulting in underreporting of these incidents (Gow et al., 2010).

A second limitation relates to the lack of a control group, so no comparisons with other patient groups were possible. To the best of our knowledge there are no studies comparing incidents and/or judicial reactions in general versus forensic psychiatric settings. Previous reviews did however reveal that overall rates of aggression were more common in forensic as compared to general psychiatric settings (Bowers et al., 2011; Linhorst & Scott, 2004). However, when solely physical aggression was considered, the difference between acute (*M* = 20.8%), general (*M* = 19.1%), and forensic (*M* = 27.9%) psychiatric settings was no longer significant (Bowers et al., 2011). Similarly, after controlling for demographic and clinical variables, Linhorst and Scott (2004) found that non-forensic patients treated in the same period in the same hospital were just as likely as forensic patients to exhibit violent behavior. Nonetheless, when comparing our study results with international studies, caution is warranted due to differences in methodology (such as a broader definition of incidents in the present study), characteristics of the population and legal systems. Different studies use highly different patient samples, different definitions of incidents or different assessments of incidents which may cause problems for comparability and generalizability. In the present study, reports were retrospectively collected on a clinical basis and not with a standardized instrument like the Staff Observation Aggression Scale-Revised (SOAS-R; Nijman et al., 1999; Palmstierna & Wistedt, 1987) or the Social Dysfunction and Aggression Scale (SDAS; Wistedt et al., 1990).

Finally, the scope of this study was limited in that it looked retrospectively at admissions. For example, the study did not provide insight in how these incidents influenced the therapeutic relationship between patients and hospital staff (Cornaggia, Beghi, Pavone, & Barale, 2011). Also, data of the three forensic units were not presented separately, so possible differences in treatment approaches could not be differentiated.

CONCLUSIONS

Despite the limitations inherent to retrospective research, some conclusions can be drawn from the present study. First, despite the negative image of forensic patients in general psychiatric facilities and in local communities, there were no incidents reported to the CPS in almost half of the population during a mean treatment period of 1.9 years. Furthermore, most of the registered incidents were not crime-related or physically violent in nature. However, incident reports had a profound impact on the treatment course. Revocation rates of conditional release were high and resulted in new periods of imprisonment. Further research is needed to determine the reasons for these high drop-out rates. It can be hypothesized that the security level of the units was not adequately matched to the risk level posed by the patients and/or the treatment was not responsive to characteristics of the patients, such as the intelligence level which was rather low. As noted in prior research, learning disabilities (particularly if undetected) may represent obstacles to engagement (Holdsworth et al., 2014). In the meanwhile, continuing investment in therapeutic relations and other strategies to promote adherence is recommended. Cornaggia et al. (2011) stressed the need for a "good ward climate", with an appropriate number of nursing staff, a non-overcrowded setting and adequate staff training. Also, keeping careful records of incidents can increase awareness of risk factors and situations, inform policy decisions regarding aggression management on the facility level and eventually prevent more serious incidents in the future (Kobes, Nijman, & Bulten, 2012). Official reconviction data clearly underestimate reoffending rates. At this moment a clear policy regarding the reporting of incidents – particularly those of an aggressive nature – during the treatment of medium security forensic psychiatric patients in Flanders is lacking. Therefore, a clear, consistent policy regarding the prosecuting of patients is recommended (Clark et al., 2012; Quanbeck, 2006).

FOOTNOTES

[1] Four missings (0.8%).

[2] If multiple offenses were present, the index offense was characterized by the most serious offense.

[3] For each internee the duration of stay was calculated by combining first admission and readmissions (if there were any). For the patients still residing in a medium security unit on December 31, 2010, admission duration was calculated to that date.

[4] Both incidents occurring on the medium security unit and during (supervised) leave from the unit were analyzed.

[5] Median two days. In eight cases the end date of the absconding was missing.

[6] Although the internment measure is executed under the authority of the CPS, any offense can basically be referred to court by the prosecutor.

[7] Person based eleven persons were involved in those incident reports.

[8] The database thus not allows to establish if these sentences resulted in an incarceration or not.

[9] A very serious violent incident was defined according to categories 8–12 (Brand, 2005).

REFERENCES

Abidin, Z., Davoren, M., Naughton, L., Gibbons, O., Nulty, A., & Kennedy, H. G. (2013). Susceptibility (risk and protective) factors for in-patient violence and self-harm: Prospective study of structured professional judgement instruments START and SAPROF, DUNDRUM-3 and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry, 13*, 197. doi: 10.1186/1471-244x-13-197

Alia-Klein, N., O'Rourke, T. M., Goldstein, R. Z., & Malaspina, D. (2007). Insight into illness and adherence to psychotropic medications are separately associated with violence severity in a forensic sample. *Aggressive Behavior, 33*(1), 86-96. doi: 10.1002/ab.20170

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Blattner, R., & Dolan, M. (2009). Outcome of high security patients admitted to a medium secure unit: The Edenfield Centre study. *Medicine, Science, and the Law, 49*(4), 247-256. doi: 10.1258/rsmmsl.49.4.247

Bowers, L., Stewart, D., Papadopoulos, C., Dack, C., Ross, J., Khanom, H., & Jeffery, D. (2011). *Inpatient violence and aggression: A literature review*. London: Kings College.

Brand, E. F. J. M. (2005). *Onderzoeksrapport PIJ-dossier 2003-C. Predictieve validiteit van de FPJ-lijst*. Den Haag: DJI.

Carr, W. A., Rotter, M., Steinbacher, M., Green, D., Dole, T., Garcia-Mansilla, A., . . . Rosenfeld, B. (2006). Structured Assessment of Correctional Adaptation (SACA): A measure of the impact of incarceration on the mentally ill in a therapeutic setting. *International Journal of Offender Therapy Comparative Criminology, 50*(5), 570-581. doi: 10.1177/0306624X06289176

Clark, C. R., McInerney, B. A., & Brown, I. (2012). The prosecution of psychiatric inpatients: Overcoming the barriers. *The Journal of Forensic Psychiatry & Psychology, 23*(3), 371-381. doi: 10.1080/14789949.2012.686622

Cornaggia, C. M., Beghi, M., Pavone, F., & Barale, F. (2011). Aggression in psychiatry wards: A systematic review. *Psychiatry Research, 189*(1), 10-20. doi: 10.1016/j.psychres.2010.12.024

Cullen, A., Jewell, A., Tully, J., Coghlan, S., Dean, K., & Fahy, T. (2015). A prospective cohort study of absconsion incidents in forensic psychiatric settings: Can we identify those at high-risk? *PLoS One, 10*(9). doi: 10.1371/journal.pone.0138819

Daffern, M., Duggan, C., Huband, N., & Thomas, S. (2008). The impact of interpersonal style on aggression and treatment noncompletion in patients with personality disorder admitted to a medium secure psychiatric unit. *Psychology, Crime & Law, 14*(6), 481-492.

Davoren, M., O'Dwyer, S., Abidin, Z., Naughton, L., Gibbons, O., Doyle, E., . . . Kennedy, H. G. (2012). Prospective in-patient cohort study of moves between levels of therapeutic security: The DUNDRUM-1 triage security, DUNDRUM-3 programme completion and DUNDRUM-4 recovery scales and the HCR-20. *BMC Psychiatry, 12*, 80. doi: 10.1186/1471-244X-12-80

Decaire, M. W., Bedard, M., Riendeau, J., & Forrest, R. (2006). Incidents in a psychiatric forensic setting: Association with patient and staff characteristics. *Canadian Journal of Nursing Research, 38*(3), 68-80.

Dinwiddie, S. H., & Briska, W. (2004). Prosecution of violent psychiatric inpatients: Theoretical and practical issues. *International Journal of Law and Psychiatry, 27*(1), 17-29. doi: 10.1016/j.ijlp.2004.01.002

Gadon, L., Johnstone, L., & Cooke, D. (2006). Situational variables and institutional violence: A systematic review of the literature. *Clinical Psychology Review, 26*(5), 515-534. doi: 10.1016/j.cpr.2006.02.002

Gow, R. L., Choo, M., Darjee, R., Gould, S., & Steele, J. (2010). A demographic study of the Orchard Clinic: Scotland's first medium secure unit. *The Journal of Forensic Psychiatry & Psychology, 21*(1), 139-155. doi: 10.1080/14789940903188956

Gradillas, V., Williams, A., Walsh, E., & Fahy, T. (2007). Do forensic psychiatric inpatient units pose a risk to local communities? *The Journal of Forensic Psychiatry & Psychology, 18*(2), 261-265. doi: 10.1080/14789940701246360

Gudjonsson, G., Rabe-Hesketh, S., & Wilson, C. (1999). Violent incidents on a medium secure unit over a 17-year period. *The Journal of Forensic Psychiatry, 10*(2), 249-263. doi: 10.1080/09585189908403679

Gudjonsson, G., Rabe-Hesketh, S., & Wilson, C. (2000). Violent incidents on a medium secure unit: The target of assault and the management of incidents. *The Journal of Forensic Psychiatry, 11*(1), 105-118. doi: 10.1080/095851800362391

Harte, J. M., Leeuwen, M. E. van, & Theuws, R. (2013). Agressie en geweld tegen hulpverleners in de psychiatrie; aard, omvang en strafrechtelijke reactie. *Tijdschrift voor Psychiatrie, 55*(5), 325-335.

Heimans, H., Beken, T. Vander, & Schipaanbord, E. (2015). Eindelijk een echte nieuwe en goede wet op de internering? Deel 1: De gerechtelijke fase. *Rechtskundig weekblad, 27*, 1043-1064. doi: 10.1080/095851800362391

Hillbrand, M. (1995). Aggression against self and aggression against others in violent psychiatric patients. *Journal of Consulting and Clininal Psychology, 63*(4), 668-671. doi: 10.1037/0022-006X.63.4.668

Holdsworth, E., Bowen, E., Brown, S., & Howat, D. (2014). Offender engagement in group programs and associations with offender characteristics and treatment factors: A review. *Aggression and Violent Behavior, 19*(2), 102-121. doi: 10.1016/j.avb.2014.01.004

Hollin, C. R., Davies, S., Duggan, C., Huband, N., McCarthy, L., & Clarke, M. (2013). Patients with a history of arson admitted to medium security: Characteristics on admission and follow-up postdischarge. *Medicine, Science, and the Law, 53*(3), 154-160. doi: 10.1258/msl.2012.012056

Jeandarme, I., Habets, P., Oei, T. I., & Bogaerts, S. (2016). Reconviction and revocation rates in Flanders after medium security treatment. *International Journal of Law and Psychiatry*. doi: 10.1016/j.ijlp.2016.02.033

Kobes, M. H., Nijman, H. H., & Bulten, E. B. (2012). Assessing aggressive behavior in forensic psychiatric patients: Validity and clinical utility of combining two instruments. *Archives of Psychiatric Nursing, 26*(6), 487-494. doi: 10.1016/j.apnu.2012.04.004

Kumar, S., Fischer, J., Ng, B., Clarke, S., & Robinson, E. (2006). Prosecuting psychiatric patients who assault staff: a New Zealand perspective. *Australasian Psychiatry*, *14*(3), 251-255. doi: 10.1111/j.1440-1665.2006.02296.x

Lamb, H. R., Weinberger, L. E., & Gross, B. H. (1999). Community treatment of severely mentally ill offenders under the jurisdiction of the criminal justice system: A review. *Psychiatric Services, 50*(7), 907-913.

Leeuwen, M. E. van, & Harte, J. M. (2011). Violence against care workers in psychiatry: Is prosecution justified? *International Journal of Law and Psychiatry, 34*(5), 317-323. doi: 10.1016/j.ijlp.2011.08.011

Linhorst, D. M., & Scott, L. P. (2004). Assaultive behavior in state psychiatric hospitals: differences between forensic and nonforensic patients. *Journal of Interpersonal Violence*, *19*(8), 857-874. doi: 10.1177/0886260504266883

Main, N., & Gudjonsson, G. H. (2006). An investigation into the factors that are associated with non-compliance in medium secure units. *The Journal of Forensic Psychiatry & Psychology, 17*(2), 181-191. doi: 10.1080/14789940500480848

McCarthy, L., & Duggan, C. (2010). Engagement in a medium secure personality disorder service: A comparative study of psychological functioning and offending outcomes. *Criminal Behaviour and Mental Health, 20*(2), 112-128. doi: 10.1002/cbm.758

McMurran, M., & Theodosi, E. (2007). Is treatment non-completion associated with increased reconviction over no treatment? *Psychology, Crime, & Law, 13*(4), 333-343. doi: 10.1080/10683160601060374

Meysman, M. (2016). *This is madness: Detention of mentally ill offenders in Europe and its impact on European cooperation*. Manuscript submitted for publication.

Muller-Isberner, J. R. (1996). Forensic psychiatric aftercare following hospital order treatment. *International Journal of Law and Psychiatry, 19*(1), 81-86.

Nijman, H. L. I., Muris, P., Merckelbach, H. L. G. J., Palmstierna, T., Wistedt, B., Vos, A. M., . . . Allertz, W. (1999). The Staff Observation Aggression Scale-Revised (SOAS-R). *Aggressive Behavior, 25*(3), 197–209.

Olver, M. E., & Wong, S. C. (2009). Therapeutic responses of psychopathic sexual offenders: Treatment attrition, therapeutic change, and long-term recidivism. *Journal of Consulting and Clinical Psychology, 77*(2), 328-336. doi: 10.1037/a0015001

Palmstierna, T., & Wistedt, B. (1987). Staff observation aggression scale, SOAS: Presentation and evaluation. *Acta Psychiatrica Scandinavica, 76*(6), 657-663.

Quanbeck, C. (2006). Forensic psychiatric aspects of inpatient violence. *The Psychiatric clinics of North America, 29*(3), 743-760. doi: 10.1016/j.psc.2006.04.011

Rotter, M., Carr, W. A., Magyar, M., & Rosenfeld, B. (2011). From incarceration to community care: Structured assessment of correctional adaptation. *Journal of the American Academy of Psychiatry and the Law, 39*(1), 72-77.

Schanda, H., Stompe, T., & Ortwein-Swoboda, G. (2009). Dangerous or merely difficult? The new population of forensic mental hospitals. *European Psychiatry, 24*(6), 365-372. doi: 10.1016/j.eurpsy.2009.07.006

Stassen, W., Habets, P., Mertens, A., Laender, J. De, & Jeandarme, I. (2014). The InReach project: From penitentiary to forensic hospital. *The International Journal of Therapeutic Communities, 35*(3), 119-126. doi: 10.1108/TC-01-2014-0002

To, W. T. (2015). *Assessment and treatment perspectives on offenders with disabilities* [Unpublished doctoral dissertation]. Ghent University, Ghent, Belgium.

Vandevelde, S., Soyez, V., Beken, T. Vander, Smet, S. De, Boers, A., & Broekaert, E. (2011). Mentally ill offenders in prison: The Belgian case. *International Journal of Law and Psychiatry, 34*(1), 71-78. doi: 10.1016/j.ijlp.2010.11.011

Wistedt, B., Rasmussen, A., Pedersen, L., Malm, U., Traskman-Bendz, L., Wakelin, J., & Bech, P. (1990). The development of an observer-scale for measuring social dysfunction and aggression. *Pharmacopsychiatry, 23*(6), 249-252. doi: 10.1055/s-2007-1014514

Wormith, J. S., & Olver, M. E. (2002). Offender treatment attrition and its relationship with risk, responsivity, and recidivism. *Criminal Justice and Behavior, 29*(4), 447-471. doi: 10.1177/0093854802029004006

Risk factors associated with inpatient violence during medium security treatment

Journal of Interpersonal Violence,
Advance online publication

Jeandarme, I., Wittouck, C., Vander Laenen, F., Pouls, C., Oei, T. I., & Bogaerts, S.

6

ABSTRACT

Violence is a common phenomenon both in regular and forensic psychiatric settings, and has a profound impact on staff and other patients. Insight into the individual risk factors associated with violence in forensic psychiatric settings is rare and is therefore the subject of this research. A retrospective file study in three medium security units in Flanders was conducted to compare non-violent inpatients with inpatients who engaged in (verbal and physical) violent behavior. Binary logistic regression analyses were used to examine which variables contributed independently to the risk of violence. The results showed that absconding during treatment was independently associated with physical violence. A personality disorder diagnosis and general non-compliance with treatment were associated with verbal violence. Both types of violence predicted early termination of treatment. Contrary to previous research, the results from the risk assessment tools were not associated with inpatient violence. Clinical implications are discussed and include, among others, that clinicians should remain vigilant for early warning signs of non-compliance during treatment.

INTRODUCTION

Violence is a common phenomenon in regular and forensic psychiatric settings, although the risk of very serious physical violence with serious injuries is limited (Cornaggia, Beghi, Pavone, & Barale, 2011; Kelly, Subica, Fulginiti, Brekke, & Novaco, 2015; Woods & Ashley, 2007). In a forensic psychiatric setting, nearly all staff reported verbal violence, 70% reported being physically assaulted during the past year, and 12% of the staff was injured so badly they had to take time off from work (Kelly et al., 2015). Besides affecting staff, patients perpetrating violence are adversely affected as well because seclusion and restraints are often used to manage aggressive behavior, and inpatient violence can result in prolonged involuntary confinement (Quanbeck, 2006). Furthermore, disruptive or violent behavior can result in discharge from treatment, although to the best of our knowledge, this has not been examined yet in forensic psychiatric settings.

Research addressing inpatient violence in forensic psychiatric settings is relatively scarce versus the amount of research conducted in civil psychiatric settings, prisons, and among professionals working in the domain of security (Bogaerts, Daalder, Van der Knaap, Kunst, & Buschman, 2008; Chan & Chow, 2014; Hogan & Ennis, 2010; Kunst, Bogaerts, & Winkel, 2009). Studying inpatient violence and victimization in forensic settings is important because one of the primary objectives of forensic psychiatric treatment is to learn violent patients how to de-escalate as well as to teach vulnerable patients, such as psychotic forensic patients, how to prevent (re)victimization. Furthermore, inpatient violence predicts violent recidivism after treatment (Spreen, Brand, Ter Horst, & Bogaerts, 2014).

Several models can explain the occurrence of inpatient violence. The internal model suggests that individual factors such as antisocial personality characteristics or impulsivity contribute to violence. The external model emphasizes the role of environmental factors such as ward design and the approach to treating staff. The situational or interactional model focuses on interactions between patients and staff members (Dickens, Piccirillo, & Alderman, 2012). Relatively few studies compared individual characteristics of violent and non-violent (forensic) patients (Bowers et al., 2011). However, knowledge about these individual inpatient risk factors is very important and can inform risk management policies and strategies, and enables staff to

recognize risk situations, thus limiting both inpatient as well as future aggressive incidents (Bowers et al., 2011).

Individual factors associated with inpatient violence in forensic psychiatric settings

First, we note that a uniform overall accepted definition of aggressive or violent behavior is lacking, which can result in different operationalizations (Klerx-Van Mierlo & Bogaerts, 2011). Some studies focused exclusively on physical violence toward others, that is, one type of interpersonal violence (Linhorst & Scott, 2004). Other studies included verbal violence in their definition of interpersonal violence and/or included other forms of violence such as self-directed violence (e.g., suicidal behavior) or violence toward objects (e.g., Chan & Chow, 2014). Here, violence refers to a specific form of interpersonal aggression that results in or threatens physical harm to another person (further interpersonal violence [IPV]). Aggression refers to a broader definition including other forms of violence such as self-directed violence or violence against property (further aggression in general [AIG]).

Two meta-analyses of the risk factors associated with inpatient AIG in forensic settings found that aggressive patients were significantly younger and—contrary to the findings in acute general psychiatric settings—more likely to be male (Dack, Ross, Papadopoulos, Stewart, & Bowers, 2013; Hogan & Ennis, 2010). However, some other single studies found no gender-aggression association (Ball, Young, Dotson, Brothers, & Robbins, 1994; Nicholls, Brink, Greaves, Lussier, & Verdun-Jones, 2009; Rogers, Watt, Gray, MacCulloch, & Gournay, 2002). Criminal history variables such as previous violent offenses or having a violent index offense were not associated with IPV in most single studies (Doyle, Dolan, & McGovern, 2002; Hoptman, Yates, Patalinjug, Wack, & Convit, 1999; Linhorst & Scott, 2004; Rasmussen & Levander, 1996; Rogers et al., 2002). One study stressed a history of violence as a predictor of AIG in general and IPV in particular (Ball et al., 1994). The number of previous psychiatric admissions was associated with IPV (Ball et al., 1994; Linhorst & Scott, 2004) and with seclusion after an AIG incident (Thomas et al., 2009).

From a clinical perspective, intellectual disability was not associated with IPV (Ball et al., 1994), but some studies found an association between low education and IPV (Harris & Varney,

1986; Hoptman et al., 1999). Personality disorders were typically associated with IPV and AIG (Chan & Chow, 2014; Gow, Choo, Darjee, Gould, & Steele, 2010), but one study failed to find an association (Ball et al., 1994). In contrast, a history of substance misuse (including alcohol misuse) at admission was not associated with IPV (Ball et al., 1994; Hoptman et al., 1999; Rogers et al., 2002; van der Kraan et al., 2014). However, a current substance-related disorder and a dual diagnosis of schizophrenia and substance misuse were associated with IPV (Chan & Chow, 2014; Hoptman et al., 1999). Psychosis or schizophrenia alone was not associated with IPV in forensic psychiatric settings (Douglas, Guy, & Hart, 2009; Hoptman et al., 1999). In general, the presence of a psychiatric diagnosis per se does not seem to be a sufficiently sensitive or reliable measure for IPV (Doyle et al., 2002), but the stage of the illness or active symptomatology might be a more valuable predictor (Gudjonsson, Rabe-Hesketh, & Wilson, 1999). For example, several studies found that positive psychotic symptoms were significantly higher in IPV and AIG (Daffern, Howell, Ogloff, & Lee, 2005; Rasmussen & Levander, 1996; Vitacco et al., 2009). Doyle and Dolan (2006) found that anger emotion regulation problems were associated with IPV, and Wang and Diamond (1999) addressed anger as a stronger predictor of IPV than previous violence. Furthermore, patients exhibiting chronic suicidal and self-injuring behaviors showed the highest levels of IPV (Hillbrand, 1995). Finally, more objective markers, such as cognitive tests or serum cholesterol may be potentially valuable predictors for IPV and AIG. However, research in this field is still in its infancy (O'Reilly et al., 2015; Sedgwick, Young, Das, & Kumari, 2016).

Other potential risk factors for violence in forensic settings relate to general misconduct during treatment. First, persistent substance misuse during treatment was associated with violent recidivism during absconding (Hildebrand, Schonberger, & Spreen, 2007). Second, low treatment engagement (defined as both attendance and active participation at core program sessions) was significantly more common in aggressive female forensic patients (Long, Dolley, & Hollin, 2012). The items unresponsive to treatment and non-compliance with remediation attempts of the Historical Clinical Risk Management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997) were both related to IPV (Abidin et al., 2013; Mudde, Nijman, van der Hulst, & van den Bout, 2011). Third, absconding was associated with AIG (Brook, Dolan, & Coorey, 1999) and with verbal but not with physical IPV (Cullen et al., 2015). Ball et al. (1994) stressed the current risk of escape as a potential indicator for IPV.

Risk assessment instruments predicting inpatient violence in forensic psychiatric settings

Besides these individual risk factors for inpatient IPV and AIG, the value of risk assessment instruments to predict inpatient violence has been studied. For example, the HCR-20 (especially the clinical scale) was useful in assessing risk for inpatient IPV and AIG in forensic psychiatric settings (Hogan & Ennis, 2010; McDermott, Edens, Quanbeck, Busse, & Scott, 2008; Mudde et al., 2011; Wilson, Desmarais, Nicholls, Hart, & Brink, 2013). Contrary to these findings, the predictive validity of the HCR-20 for IPV in forensic units was poor in field validity studies (Jeandarme, Pouls, De Laender, Oei, & Bogaerts, 2016; Neal, Miller, & Shealy, 2015; Pedersen, Ramussen, & Elsass, 2012; Vojt, Thomson, & Marshall, 2013). Along with the HCR-20, Hogan and Ennis (2010) found that the Hare psychopathy scales were predictive of AIG. Other meta-analyses investigated the association between the Hare psychopathy scales and institutional violence in different settings among which forensic psychiatric settings. They found a more robust association with general inpatient misconduct compared with physical violence and with Psychopathy Checklist–Revised (PCL-R) Factor 2 compared with PCL-R Factor 1 (Guy, Edens, Anthony, & Douglas, 2005; Walters, 2003). Furthermore, psychopathy was a stronger predictor for proactive or instrumental IPV than for reactive IPV in a forensic psychiatric hospital (Vitacco et al., 2009). The Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice, & Cormier, 2006) is another risk-assessment instrument designed to predict violence. Some studies found that the VRAG was predictive of inpatient IPV and AIG (Doyle et al., 2002; McDermott, Dualan, & Scott, 2011; Snowden, Gray, Taylor, & Fitzgerald, 2009), although others found no effect (Chu, Thomas, Ogloff, & Daffern, 2013; McDermott, Quanbeck, Busse, Yastro, & Scott, 2008). In Flanders, the VRAG predicted institutional IPV only marginally in forensic psychiatric offenders with an intellectual disability (Pouls, Jeandarme, & Habets, 2014). Unlike these previous studies, other schemes such as the Broset Violence Checklist (BVC; Almvik, Woods, & Rasmussen, 2000) or the Dynamic Appraisal of Situational Aggression (DASA; Chan & Chow, 2014; Ogloff & Daffern, 2006) have been developed specifically to assess the risk for inpatient violence, but these are used infrequently in forensic settings (for example, see, Woods, Olver, & Muller, 2015).

Methodological problems associated with studies on inpatient violence

The comparison of studies investigating inpatient violence and aggression is hampered by several methodological problems. First, as mentioned earlier, different definitions and operationalizations of violent or aggressive behavior are used and may contribute to some of the contradictory findings. A literature review on psychiatric inpatient aggression identified the most common operationalizations: (a) interpersonal physical violence, (b) a combination of interpersonal physical and verbal violence and aggression toward objects, and (c) a combination of interpersonal physical and verbal violence, aggression toward objects, and self-directed violence (Bowers et al., 2011).

Second, the review of Dack et al. (2013) shows that the studies differ in their method of data collection. The majority of studies used patient notes or standard incident reports, while some used standardized and validated scales designed to measure aggression, such as the Staff Observation Aggression Scale–Revised (SOAS-R; Nijman et al., 1999). In addition, it is likely that the terms inpatient and institutional are used interchangeably. Inpatient violence can occur on different occasions: in the unit, in the hospital premises, or during (un)supervised leave outside of the hospital premises. Institutional violence only refers to violence on the ward. These specifications are rarely reported.

Third, studies use different follow-up periods, whereby higher base rates for interpersonal violence can be expected with longer follow-up periods. In addition, different statistical analyses that do or do not control or adjust for potential confounding factors may result in different interpretations.

Present study

This study contributes to the research on IPV within a forensic psychiatric setting by examining risk factors for IPV during medium security treatment. Prior research rarely evaluates separate forms of violence. Therefore, we will analyze several types of IPV (verbal, physical, and a combination of both). This study also analyzes whether IPV predicts treatment dropout. The following research questions are addressed:

- *Research Question 1.* How often does IPV occur during forensic psychiatric treatment and what is the nature of the IPV?
- *Research Question 2.* Which individual demographic, clinical, and criminogenic risk factors for inpatient violence—identified in previous research—differentiate between several levels of IPV?
- *Research Question 3.* Which risk factors are independently associated with the occurrence of verbal, physical, or combined IPV?
- *Research Question 4.* Does the occurrence of verbal, physical, or combined IPV predict dropout from treatment?

MATERIAL AND METHOD

Setting

The study was conducted at the three forensic medium security units (MSUs) in Flanders (the Dutch-speaking part of Belgium). These units are part of the general psychiatric hospitals and are located in the communities of Bierbeek, Rekem, and Zelzate. All patients were found not guilty by reason of insanity (NGRI, in Belgium referred to as internees) after having committed an offense. They were submitted to an internment measure to protect society from further offenses. Conditional release from an internment measure is linked to mandatory treatment under the supervision of a regional court—the Commission of the Protection of Society (CPS). Violent (and other) incidents occurring during the treatment are reported to the CPS on a regular basis by the MSUs and/or the probation officer. In the case of incidents, the CPS may interrupt treatment due to revocation of conditional release. This would typically be the case when either the MSU or the internee (less common) decide to stop the treatment. In case of very serious incidents, i.e., when the safety of the public can no longer be guaranteed by the MSU, the treatment can be stopped. Because no high security units were established at the time of the study, revocation of conditional release resulted in incarceration either for a time-out period or for a longer confinement.

The MSUs were implemented in Flanders in 2001 with the goal of providing a treatment setting for NGRI patients who do not require very secure hospital care, but who are deemed too

dangerous or unsuitable for a general psychiatric ward or outpatient care. MSUs typically admit psychotic as well as personality-disordered internees. Internees with a primary diagnosis of substance use disorder, intellectual disability, or paraphilia are excluded from treatment, although all these conditions can be present as co-morbid disorders (see Boers, Vandevelde, Soyez, De Smet, & Ting to, 2011, for a description of treatment programs).

Study Population

The population (*N* = 531) was predominantly male (94.9%, *n* = 504) with a Belgian nationality (90.1%, *n* = 475). The mean age at first admission to a MSU was 36.5 years (*SD* = 10.82 years, range = 18.8–73.4 years). The average length of stay in these units was 676.4 days (*SD* = 505.90 days, range = 8–2,729 days). About 10% of the internees (*n* = 65) were still in treatment at the end of the study. Of the remaining group (*n* = 461), a third (31.7%, *n* = 146) dropped out prematurely from treatment.

Participants were found NGRI for a variety of offenses, where the most serious offense consisted of violence (77.2%, *n* = 410), property crimes (18.6%, n = 99), drug-related crimes (2.3%, *n* = 12), sexual non-violent crimes (0.8%, *n* = 4), and other offenses such as destruction of property (1.1%, *n* = 6). Prior contact with the criminal justice system was common; only a minority (15.6%, *n* = 83) were registered as first offender. On average, patients had been sentenced on 6.3 occasions (*SD* = 5.73, range = 1–40) for a general offense and on 2.4 occasions (*SD* = 1.91, range = 0–11) for a violent offense.

The most common diagnoses according to the Diagnostic and Statistical Manual of mental disorders-IV-Text Revision (DSM-IV-TR; American Psychiatric Association, 2000) were personality disorders (70.6%, *n* = 375), substance use disorders (56.7%, *n* = 301), psychotic disorders (43.9%, *n* = 233), and intellectual disability (12.8%, *n* = 68). The mean total score on the PCL-R (Hare, 2003; *n* = 224) was 21.1 (*SD* = 6.58, range = 6–36.8). The mean HCR-20 score (*n* = 235) was 24.5 (*SD* = 5.34, range = 10.5–36), and the mean VRAG score (*n* = 230) was 8 (*SD* = 10.94, range = −26 to +38).

Procedure and outcome measures

Data were collected for all patients treated between 2001 and 2010. All patients were asked for passive consent. Eleven patients refused, leaving a total study population of 531 patients. This

comprised approximately the whole Flemish medium security population over a 10-year study period.

Violent incidents were registered based on the incident reports found in CPS files. Person-based information was gathered through CPS as well as through hospital files. Information on treatment variables was limited to information that was reported to the CPS. This study focused on IPV, which referred to the intentional use of physical force or power—threatened, attempted, or actual—against another person. Verbal IPV referred to threatened violence; physical IPV referred to attempted or actual IPV. Four groups were constructed with increasing severity of violence: (a) patients with no violent incidents, (b) patients with only verbally violent incidents, (c) patients with only physically violent incidents, and (d) patients with both verbally and physically violent incidents.

Non-compliance referred to not adhering to treatment rules, for example, not engaging in treatment plans, refusing prescribed medication, or drinking alcohol during treatment, whereas non-compliance alcohol was limited to drinking alcohol. Drug use referred to the use of illegal substances or nonprescription medication as well as to dealing of illicit substances. Absconding referred to absconding from (un)supervised leave or absconding without permission to leave the unit or premises. Dropout referred to treatment dropout. As explained earlier, internees whose treatment was stopped were incarcerated again.

The VRAG was scored on a convenience sample of 230 patients; other risk assessment scores were field validity scores. In case of multiple scores, the first assessment was selected.

The study was approved by the Ethics committee of Antwerp University Hospital.

Data analysis

SPSS Version 22 was used for the statistical analyses. Patient base rates (violent patients/sample × 100) and event base rates (incidents/sample × 100) were calculated (based on Bowers et al., 2011). Differences between the four patient groups were tested using the chi-square or Fisher exact test in case of categorical variables. The means between the patient groups were compared with between-group one-way ANOVAs if needed after root square log transformation. Post hoc comparisons used Tukey HSD or Dunnett T3 in case of unequal variances. Effect sizes were measured using Cramer's *V* for categorical variables and partial eta squared and r for continuous

variables. McNemar's test was used on paired nominal data. Significant bivariate associations were entered in multiple logistic regression analyses. An additional regression analysis investigated whether IPV would predict dropout. The correlation between length of stay and the occurrence of (physically and verbally) violent incidents was analyzed using point-biserial correlations to determine whether this effect needed to be controlled for in subsequent analyses.

Some analyses were based on smaller samples because the data sources were characterized by some missing data; this is noted throughout the article and in Table 2.

RESULTS

Prevalence of violence

Between 2001 and 2010, IPV incidents were recorded for 21.7% of the admitted patients (n = 115). In total, 189 IPV incidents were reported (event base rate = 35.6%). The patient base rate was 13.7% (n = 73) for verbal IPV and 13% (n = 69) for physical IPV. The event base rate for verbal IPV was 17.5% and 18.1% for physical IPV. Table 1 shows the nature of the incidents in more detail. Twenty-seven patients exhibited both verbal and physical IPV. Neither type of IPV were associated (McNemar, p = .75).

Most patients involved in incidents were violent in the unit or hospital premises (82.6%, n = 95); a minority were violent only during (un)supervised leave (7%, n = 8) or on both occasions (10.4%, n = 12).

The relationship between patients' length of stay and the occurrence of violent incidents (r = .02, p = .59), physically violent incidents (r = .07, p = .12), and verbally violent incidents (r = −.02, p = .72), was not significant.

Variables contributing to the risk of violence

Four groups of patients were compared in Table 2: (a) patients with no IPV incidents (n = 416), (b) patients with only verbal IPV incidents (n = 46), (c) patients with only physical IPV incidents (n = 42), and (d) patients with both verbal and physical IPV incidents (n = 27).

Table 1. Nature of IPV (event- and person-related)

	Number of patients involved *n* (%)	Number of incidents *n* (%)
Any IPV	115 (21.7%)	189 (100%)
Verbal IPV	73 (13.7%)	93 (49.2%)
Physical IPV	69 (13%)	96 (50.8%)
Other violence,for example, assault	54	72
Arson in an occupied property	9	9
Violent property offenses, for example, snatching	6	8
Sexual hands-on offense	5	5
Attempted manslaughter	2	2

Note. IPV = interpersonal violence

Demographic, criminal history, and clinical variables

There was a marginal significant negative effect of age on severity of IPV, $F(3, 527) = 2.85$, $p = .04$. There was also a significant linear trend indicating that as the age at admission decreased, the severity of IPV increased proportionally, $F(1, 527) = 8.52$ $p = .00$. The magnitude in the differences of the means and effect size was small (partial eta squared = .02). Post hoc comparisons to evaluate pairwise differences among group means were conducted with the Tukey HSD test because equal variances were tenable. Tests revealed significant pairwise differences between the mean age of the group with no IPV ($M = 36.9$, $SD = 10.90$, CI = [35.81, 37.91]) compared with the combined verbally and physically IPV group ($M = 31.0$, $SD = 8.43$, CI = [27.62, 34.29]), $p = .03$. The verbal and physical IPV groups did not significantly differ from the other two groups.

No significant differences were found among the groups regarding criminal justice involvement.

In general, personality disorders were associated with the severity of IPV. There was a significant difference between the percentage of patients with a personality disorder in the group with no IPV (67.5%), verbal IPV (89.1%), physical IPV (71.4%), and the combined IPV group (85.2%), $\chi^2(3) = 12.26$, $p = .01$, Cramer's $V = .15$.

Table 2. Person-related characteristics of (verbal and physical) IPV

		No IPV (n = 416)		Verbal IPV (n = 46)		Physical IPV (n = 42)		Verbal and physical IPV (n = 27)			
Categorical variables	*N*	*n*	%	*n*	%	*n*	%	*n*	%	*p*	*Cramer's V*
Demographic variables											
Female	531	21	5.0	2	4.3	3	7.1	1	3.7	.90	
Belgian nationality	527	376	91.3	39	84.8	36	85.7	24	88.9	.31	
Previous psychiatric admissions	521	328	80.6	38	82.6	34	82.9	25	92.6	.52	
Criminal justice involvement											
Index offense violence	531	312	75.0	36	78.3	37	88.1	25	92.6	.05	
Treatment characteristics											
Suicide attempts	531	15	3.6	0	0	1	2.4	2	7.4	.32	
Drug use/traffic	531	48	11.5	5	10.9	9	21.4	7	25.9	.06	
Absconding	531	110	26.4	13	28.3	22	52.4	13	48.1	.00***	.18
Non-compliance	531	92	22.1	29	63.0	23	54.8	14	51.9	.00***	.32
Non-compliance alcohol	531	38	9.1	11	23.9	14	33.3	8	29.6	.00***	.24
Clinical characteristics											
Psychiatric diagnosis											
Psychotic disorder	531	188	45.2	16	34.8	17	40.5	12	44.4	.57	
Substance use disorder	531	240	57.7	26	56.5	20	47.6	15	55.6	.66	
Intellectual disability	531	51	12.3	5	10.9	9	21.4	3	11.1	.39	
Personality disorder	531	281	67.5	41	89.1	30	71.4	23	85.2	.01**	.15
Antisocial personality disorder	531	95	22.8	15	32.6	12	28.6	10	37.0	.18	

Table 2. Person-related characteristics of (verbal and physical) interpersonal violence (continued)

		No IPV (n = 416)		Verbal IPV (n = 46)		Physical IPV (n = 42)		Verbal and physical IPV (n = 27)			
Continuous variables	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*p*	*r*
Demographic variables											
Age at first admission	531	36.9	10.90	37.3	10.18	35.2	11.28	31.0	8.43	.04*	.13
Criminal justice involvement											
Age at first sentence	531	25.0	9.29	25.2	10.51	23.7	7.45	22.9	5.87	.57	
Average sentences index included	531	6.4	5.80	5.2	4.60	6.5	7.07	5.6	3.76	.52	
Average sentences index included violence	531	2.4	1.97	2.4	1.82	2.5	1.71	2.3	1.35	.96	
Clinical characteristics											
PCL-R Total	224	20.8	6.65	22.4	5.84	23.1	5.02	20.0	7.83	.30	
PCL-R Factor 1	215	8.6	3.40	9.2	3.05	8.5	3.24	8.0	4.14	.76	
PCL-R Factor 2	213	10.0	4.05	11.6	2.95	11.7	2.74	10.8	3.41	.15	
HCR-20 Total	244	24.1	5.33	26.5	4.49	27.0	4.18	24.6	5.81	.05	
HCR-20 H-scale	244	13.9	3.30	15.0	3.20	15.7	1.90	14.2	2.84	.11	
HCR-20 C-scale	244	4.7	1.81	5.2	1.76	5.2	1.82	4.5	1.90	.38	
HCR-20 R-scale	244	5.7	2.08	6.3	1.77	6.3	2.02	6.0	2.00	.37	
VRAG Total	230	7.7	11.14	8.2	8.98	6.8	10.53	13.1	11.74	.40	

Note. IPV = interpersonal violence; PCL-R = Psychopathy Checklist-Revised; HCR-20 = Historical, Clinical, Risk management-20, VRAG = Violence Risk Appraisal Guide

* $p < .05$. ** $p < .01$. *** $p < .001$.

Treatment characteristics

Several treatment characteristics were associated with the occurrence and severity of IPV. There was a significant difference between the percentage of absconders in the group with no IPV (26.4%), verbal IPV (28.3%), physical IPV (52.4%), and combined IPV group (48.1%), $\chi^2(3) = 16.89$, $p = .00$, Cramer's $V = .18$. There was also a significant difference between the percentage of non-compliers in the group with no IPV (22.1%), verbal IPV (63.0%), physical IPV (54.8%), and the combined IPV group (51.9%), $\chi^2(3) = 54.88$, $p = .00$, Cramer's $V = .32$. Violating the alcohol prohibition was associated with the severity of IPV. There was a significant difference between the percentage of patients using alcohol in the group with no IPV (9.1%), verbal IPV (23.9%), physical IPV (33.3%), and the combined IPV group (29.6%), Fisher exact, $p = .00$, Cramer's $V = .24$. Finally, dropout was associated with the severity of IPV. There was a significant difference between the percentage of patients who dropped out from treatment in the group with no IPV (22.3%), verbal IPV (60.0%), physical IPV (66.7%), and the combined IPV group (66.7%), $\chi^2(3) = 64.87$, $p = .00$, Cramer's $V = .38$.

Logistic Regression Analyses

Variables that showed a significant difference on the bivariate level (absconding, non-compliance, non-compliance alcohol, personality disorder, and age at first admission) for severity level of IPV were entered in the logistic regression models to examine which variables contributed independently to the risk of inpatient verbal, physical, and combined verbal/physical IPV. In these analyses, the dependent variables were (a) verbal IPV ($n = 46$), (b) physical IPV ($n = 42$), and (c) verbal in combination with physical IPV ($n = 27$). The enter option was used for all analyses, and the results can be found in Table 3.

The first binary logistic regression analysis examined the association with verbal IPV and included 531 of the 531 internees. The model was statistically significant, $\chi^2(5) = 33.14$, $p = .00$, and correctly classified 91.3% of the cases. A Nagelkerke's R^2 of .14 showed that the global model explained 14% of the variation in the dependent variable. Internees who did not comply were six times more likely to be verbally violent and internees with a personality disorder were three times more likely to be verbally violent.

The second binary logistic regression analysis, which examined the association with physical IPV, included 531 of the 531 internees. The model was statistically significant, $\chi^2(5) = 19.87$, $p = .00$, and correctly classified 92.1% of the cases. A Nagelkerke's R^2 of .09 showed that the global model explained 9% of the variation in the dependent variable. Internees who absconded were twice as likely to be physically violent.

The third binary logistic regression analysis, which examined the association with combined verbal and physical IPV, included 531 of the 531 internees. The model was statistically significant, $\chi^2(5) = 17.16$, $p = .00$, and correctly classified 94.9% of the cases. A Nagelkerke's R^2 of .10 showed that the global model explained 10% of the variation in the dependent variable. Younger internees were more likely to be both verbally and physically violent.

Prediction of Treatment Dropout

Finally, we studied whether IPV would predict treatment dropout. For these analyses, the dependent variable was dropout from treatment and the independent variables were verbal IPV, physical IPV, and combined verbal and physical IPV. The enter option was used for all analyses. The binary logistic regression analysis included 461 of the 531 internees. The model was statistically significant, $\chi^2(5) = 61.24$, $p = .00$, and correctly classified 74.6% of the cases. A Nagelkerke's R^2 of .17 showed that the global model explained 17% of the variation in the dependent variable. Verbal violence predicted dropout from treatment (β = 1.65, Wald = 22.67, *df* = 1, $p < .00$, Exp(β) = 5.21, CI = [2.64, 10.29]), as well as physical violence (β = 1.94, Wald = 28.59, *df* = 1, $p < .00$, Exp(β) = 6.95, CI = [3.42, 14.15]) and combined verbal and physical violence (β = 1.94, Wald = 18.46, *df* = 1, $p < .00$, Exp(β) = 6.95, CI = [2.87, 16.83]).

DISCUSSION

The aim of the present study was to shed light on the prevalence of inpatient IPV in the three medium security settings in Flanders. Verbal, physical, and both verbal and physical IPV patients were compared with patients with no IPV via several sociodemographic, clinical, and criminogenic factors. This study contributes to the limited research in the area of inpatient violence in forensic psychiatry (Dack et al., 2013). The study retrospectively investigated violent behavior through the

analysis of incidents reported in judicial files. The study also analyzed whether IPV would predict dropout from treatment. The person and event base rate for physical IPV was low versus the base rates found in a meta-analysis of studies in forensic psychiatric settings (Bowers et al., 2011). When controlling for other variables, distinct associations were found for different forms of IPV. Overall, few risk factors for inpatient IPV were found. This is not surprising because the study mainly focused on static individual patient factors (Dack et al., 2013). More surprisingly, common risk assessment instruments such as the HCR-20, PCL-R, and VRAG were not associated with inpatient IPV. All degrees of IPV predicted dropout from treatment.

Only one *demographic* factor differentiated between the severity of violence. In line with the meta-analysis of Dack et al. (2013), IPV patients were significantly younger than non-violent patients. Also, the severity of violence increased as the age of admission decreased. After controlling for other variables, age was only associated with the combination of both verbal and physical violence. Previous research found that in forensic settings, women show significantly more violent behavior than men (Dack et al., 2013), although other studies did not (Daffern et al., 2005; Nicholls et al., 2009). We found no gender differences similar to Daffern et al. (2005); this might be due to the limited number of female patients.

With respect to *criminal justice involvement*, having a violent index offense or a violent history was not associated with IPV. This was in line with other research (Doyle et al., 2002; Linhorst & Scott, 2004). These studies showed that past violence is a poor predictor of inpatient IPV in samples with high numbers of previous violence.

More associations were found between *treatment variables* and violence. However, suicide attempts while in treatment were not associated with the severity of violence, which is in line with Hillbrand (1995). He found that having a history of suicide attempts alone was not associated with violence, although severe and chronic self-injurious and suicidal behavior did predict violence. Other variables associated with the severity of IPV related to treatment- interfering factors more generally (non-compliance, non-compliance alcohol, and absconding). After controlling for other variables, non-compliance was only associated with verbal IPV. Our definition of non-compliance (a combination of non-engagement in treatment and not adhering to treatment rules) was broader than the definitions found in other forensic studies, but added to the research and stated that non-cooperative and non-engaged patients were more likely to be

violent (Long et al., 2012; Mudde et al., 2011). In line with Ball et al. (1994), absconding was associated with the severity of IPV, but this effect disappeared after controlling for other variables for verbal violence and the combined outcome measure; it remained significant for physical violence. This study somewhat contradicts Cullen et al. (2015), who found an association with verbal but not with physical IPV.

Alcohol use was associated with the severity of IPV, while illicit drug use was not. After controlling for other variables, alcohol use was no longer associated with IPV. A few studies investigated the influence of ongoing substance use during treatment, but the relationship between substance use and crime is still poorly understood. Several factors can mediate or moderate this relationship such as the severity of the substance-related disorder; individual psychological, social, and neurobiological characteristics; situational factors; and the expectations regarding the psychopharmacological effects of a particular substance (Lammers et al., 2014). As van der Kraan et al. (2014) mentioned, it is relevant to identify whether substance misuse is a primary or a secondary criminogenic need.

The severity of IPV also showed associations with one *clinical variable*—a diagnosis of personality disorder. This was consistent with other research (Chan & Chow, 2014). After controlling for other variables, a personality disorder diagnosis was only associated with verbal violence. Whittington and Richter (2006) argued that the practical relevance of psychopathological and personality factors for dealing with aggressive patients remains unproven. In the present study, psychopathology was not associated with the severity of violence. More specifically, a diagnosis of psychosis was not associated with the severity of violence, thereby confirming the meta-analytic work by Douglas et al. (2009). A possible explanation for these findings is that the stage of the illness or the presence of active symptoms might be a more valuable predictor of violence than diagnosis as such (Daffern et al., 2005).

More surprisingly, this study found no association between the PCL-R, HCR-20, and VRAG and severity of inpatient IPV. Although previous studies found mixed results (Chu et al., 2013; Doyle et al., 2002; Endrass, Rossegger, Frischknecht, Noll, & Urbaniok, 2008; Snowden et al., 2009), the VRAG was not associated with severity of violence here. Also, in contrast to other research (McDermott, Edens, et al., 2008; Mudde et al., 2011), the HCR total and subscales were not associated with the severity of IPV. One of the explanations for this finding could be that the

HCR-20 scores in this study were field validity scores, which are poor to even non-significant predictors of inpatient violence (Jeandarme et al., 2016; Neal et al., 2015; Pedersen et al., 2012; Vojt et al., 2013). In the literature, the association of psychopathy with institutional violence is well established, although some studies found no association (e.g., Rasmussen & Levander, 1996) or showed a more nuanced picture. In this study, psychopathy was not associated with the severity of violence. It may well be that in forensic psychiatric settings, the predictive accuracy of risk assessment schemes was reduced due to risk management strategies. These are implemented when a high risk patient was identified.

All forms of violence (verbal, physical, and a combination of both) predicted dropout from treatment. Two thirds of the patients with physical violence also dropped out from treatment. The present study did not investigate whether the violent incidents actually caused the expulsion from treatment, but it did highlight that inpatient IPV was not only associated with prolonged confinement as stated by Quanbeck (2006) but also associated with early treatment termination. The association of violence with a high risk of treatment dropout raises concern because a meta-analysis investigating cognitive– behavioral offender treatment programs has shown that offenders who do not complete treatment have higher recidivism rates than offenders who complete the treatment program. Furthermore, non-completers had higher recidivism rates than those who were not offered treatment (d = −.16; McMurran & Theodosi, 2007). While this study did not investigate this, it is plausible that being violent was a reason for treatment expulsion (Long et al., 2012).

Limitations

Some limitations of the current study must be understood to interpret the results appropriately. First, the outcome measure of the study was limited to violence against others or IPV whereas other studies examined a much wider range of violent or aggressive behavior including aggression to property or self-directed violence. As such, methodological issues hamper the comparability of the current results with other studies.

Second, violent incidents were selected from incident reports recorded in judicial files, which tend to underestimate the number of assaults. Because there is neither a formal obligation nor standardized reporting protocol in Flemish forensic institutions to the CPS court, the dark

number of violent incidents remains unknown. Some studies used a standardized instrument such as the SOAS-R (Nijman et al., 1999; Palmstierna & Wistedt, 1987), while other studies used a range of approaches. This also contributes to methodological problems in comparing study results.

Third, severity was defined according to whether the IPV was verbal or physical (or combined). This may be somewhat arbitrary if, for example, a death threat (verbal) is compared with a push (physical).

Fourth, data of the three forensic units were not presented separately, and thus, possible differences in treatment and reporting approaches could not be differentiated.

Fifth, with respect to the bivariate analyses, the hypotheses were based on the unadjusted p values with a probability that some of the true null hypotheses were rejected.

Finally, the study was retrospective, which limits its utility. We focused on individual risk factors for interpersonal violence and did not analyze protective factors. Also, it was, for example, not possible to gain insight into how these incidents influenced the therapeutic relationship between patients and hospital staff and what the role of patient–staff interactions might have been on the occurrence of violence (Cornaggia et al., 2011). In this respect, this study is only the first step toward the management and prevention of violent incidents. It would be very interesting to study de-escalation techniques applied by staff in different settings, and to examine whether the application of different techniques can impact the prevalence and severity of violent incidents (Hallett & Dickens, 2015).

CLINICAL IMPLICATIONS AND CONCLUSION

The most surprising finding regarding IPV was perhaps that in MSUs where the vast majority of the patients had a history of violence, most patients did not behave violently during their stay. This low figure possibly reflects adequate risk management of staff but may also be explained by underreporting. Accurate reporting is important both to enhance risk management and to limit liability issues. This study showed that IPV patients misbehaved more generally during treatment, that is, more violent patients absconded, did not comply with hospital rules, and kept on drinking alcohol during treatment even though this was prohibited. In fact, after controlling for other

variables, treatment characteristics were the only characteristics that predicted IPV. The commonly used risk assessment instruments and other well-established risk factors for community violence such as previous violence were not useful predictors. Patients who are not actively collaborating with staff or who do not comply with hospital rules are more prone to become engaged in coercive interactions that may escalate into IPV. Two thirds of the patients with physical IPV dropped out from treatment versus one in five patients without IPV. This increased risk for dropout was not exclusively related to physical violence—60% of the patients with verbal IPV dropped out from treatment. Thus, it is important to be vigilant for verbal violence and to investigate this form of violence further. Although not studied here, threats may precede physical violence and may be a powerful predictor of violent acts in forensic patients as already noted by Ball et al. (1994) and Woods et al. (2015).

In sum, these results suggest that dynamic treatment variables can act as warning signals for IPV. As was noted in a review on patients' perspectives on violence, patients stated that they had often given staff warnings about potential violence or showed clear signals of distress over a long period, without receiving adequate intervention. This causes misunderstanding, and feelings of being ignored (Gudde, Olso, Whittington, & Vatne, 2015). Treatment staff should carefully monitor incidents, for instance, by using an instrument to monitor treatment progress and incidents such as the Dutch Instrument for Forensic Treatment Evaluation (Schuringa, Spreen, & Bogaerts, 2014).

Furthermore, the need for a more dynamic and contextual approach in investigating predictors of inpatient IPV was highlighted by the few individual static risk factors showing different associations with verbal and physical IPV. Indeed, triggers for violence are multifactorial and involve interplay between individual, situational, and structural factors (Dickens et al., 2012). Cornaggia et al. (2011) noted in this respect that there is a particular need for an appropriate number of nurses, a non-overcrowded setting, nurses' training, and a "good warm climate" in dealing with violence (p. 18). In other words, a narrow focus on individual patients is only part of the solution to reduce inpatient violence. As stated by Bader and Evans (2015), we cannot expect severely ill patients to recover "in polluted, toxic environments" (p. 181).

REFERENCES

Abidin, Z., Davoren, M., Naughton, L., Gibbons, O., Nulty, A., & Kennedy, H. G. (2013). Susceptibility (risk and protective) factors for in-patient violence and self-harm: Prospective study of structured professional judgement instruments START and SAPROF, DUNDRUM-3 and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry, 13*, 197. doi: 10.1186/1471-244x-13-197

Almvik, R., Woods, P., & Rasmussen, K. (2000). The Broset violence checklist: Sensitivity, specificity, and interrater reliability. *Journal of Interpersonal Violence, 15*(12), 1284-1296. doi: 10.1177/088626000015012003

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Bader, S. M., & Evans, S. E. (2015). Implementing an ecological approach to violence reduction at a forensic psychiatric hospital: Approaches and lessons learned. *CNS Spectrums, 20*(3), 177-181. doi: 10.1017/S1092852915000176

Ball, E. M., Young, D., Dotson, L. A., Brothers, L. T., & Robbins, D. (1994). Factors associated with dangerous behavior in forensic inpatients: Results from a pilot study. *Bulletin of the American Academy of Psychiatry and the Law, 22*(4), 605-620.

Boers, A., Vandevelde, S., Soyez, V., Smet, S. De, & To, W.T. (2011). Het zorgaanbod voor geïnterneerden in België. *Panopticon, 2*, 17-38.

Bogaerts, S., Daalder, A. L., Knaap, L. M. Van der, Kunst, M. J., & Buschman, J. (2008). Critical incident, adult attachment style and posttraumatic stress disorder: A comparison of three groups of security workers. *Social Behavior and Personality, 36*(8), 1063-1072. doi: 10.2224/sbp.2008.36.8.1063

Bowers, L., Stewart, D., Papadopoulos, C., Dack, C., Ross, J., Khanom, H., & Jeffery, D. (2011). *Inpatient violence and aggression: A literature review*. London: Kings College.

Brook, R., Dolan, M., & Coorey, P. (1999). Absconding of patients detained in an English special hospital. *The Journal of Forensic Psychiatry, 10*(1), 46-58. doi: 10.1080/09585189908402138

Chan, O., & Chow, K. K. (2014). Assessment and determinants of aggression in a forensic psychiatric institution in Hong Kong, China. *Psychiatry Research, 220*(1-2), 623-630. doi: 10.1016/j.psychres.2014.08.008

Chu, C. M., Thomas, S. D., Ogloff, J. R., & Daffern, M. (2013). The short- to medium-term predictive accuracy of static and dynamic risk assessment measures in a secure forensic hospital. *Assessment, 20*(2), 230-241. doi: 10.1177/1073191111418298

Cornaggia, C. M., Beghi, M., Pavone, F., & Barale, F. (2011). Aggression in psychiatry wards: A systematic review. *Psychiatry Research, 189*(1), 10-20. doi: 10.1016/j.psychres.2010.12.024

Cullen, A. E., Jewell, A., Tully, J., Coghlan, S., Dean, K., & Fahy, T. (2015). A prospective cohort study of absconsion incidents in forensic psychiatric settings: Can we identify those at high-risk? *PLoS One, 10*(9), e0138819. doi: 10.1371/journal.pone.0138819

Dack, C., Ross, J., Papadopoulos, C., Stewart, D., & Bowers, L. (2013). A review and meta-analysis of the patient factors associated with psychiatric in-patient aggression. *Acta Psychiatrica Scandinavica, 127*(4), 255-268. doi: 10.1111/acps.12053

Daffern, M., Howell, K., Ogloff, J., & Lee, J. (2005). Individual characteristics predisposing patients to aggression in a forensic psychiatric hospital. *The Journal of Forensic Psychiatry & Psychology, 16*(4), 729 – 746. doi: 10.1080/14789940500345595

Dickens, G., Piccirillo, M., & Alderman, N. (2012). Causes and management of aggression and violence in a forensic mental health service: Perspectives of nurses and patients. *International Journal of Mental Health Nursing, 22*(6), 532-544. doi: 10.1111/j.1447-0349.2012.00888

Douglas, K. S., Guy, L. S., & Hart, S. D. (2009). Psychosis as a risk factor for violence to others: A meta-analysis. *Psychological Bulletin, 135*(5), 679-706. doi: 10.1037/a0016311

Doyle, M., & Dolan, M. (2006). Evaluating the validity of anger regulation problems, interpersonal style, and disturbed mental state for predicting inpatient violence. *Behavioral Sciences & the Law, 24*(6), 783-798. doi: 10.1002/bsl.739

Doyle, M., Dolan, M., & McGovern, J. (2002). The validity of North American risk assessment tools in predicting in-patient violent behaviour in England. *Legal and Criminological Psychology, 7*(2), 141–154. doi: 10.1348/135532502760274756

Endrass, J., Rossegger, A., Frischknecht, A., Noll, T., & Urbaniok, F. (2008). Using the Violence Risk Appraisal Guide (VRAG) to predict in-prison aggressive behavior in a Swiss offender population. *International Journal of Offender Therapy and Comparative Criminology, 52*(1), 81-89. doi: 10.1177/0306624X07301643

Gow, R. L., Choo, M., Darjee, R., Gould, S., & Steele, J. (2010). A demographic study of the Orchard Clinic: Scotland's first medium secure unit. *The Journal of Forensic Psychiatry & Psychology, 21*(1), 139-155. doi: 10.1080/14789940903188956

Gudde, C. B., Olso, T. M., Whittington, R., & Vatne, S. (2015). Service users' experiences and views of aggressive situations in mental health care: A systematic review and thematic synthesis of qualitative studies. *Journal of Multidisciplinary Healthcare, 8*, 449-462. doi: 10.2147/JMDH.S89486

Gudjonsson, G., Rabe-Hesketh, S., & Wilson, C. (1999). Violent incidents on a medium secure unit over a 17-year period. *The Journal of Forensic Psychiatry, 10*(2), 249-263. doi: 10.1080/09585189908403679

Guy, L. S., Edens, J. F., Anthony, C., & Douglas, K. S. (2005). Does psychopathy predict institutional misconduct among adults? A meta-analytic investigation. *Journal of Consulting and Clinical Psychology, 73*(6), 1056-1064. doi: 10.1037/0022-006X.73.6.1056

Hallett, N., & Dickens, G. L. (2015). De-escalation: A survey of clinical staff in a secure mental health inpatient service. *International Journal of Mental Health Nursing, 24*(4), 324-333. doi: 10.1111/inm.12136

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist*. Toronto, ON: Multi-Health Systems.

Harris, G., & Varney, G. (1986). A ten-year study of assaults and assaulters on a maximum security psychiatric unit. *Journal of Interpersonal Violence, 1*(2), 173-191.

Hildebrand, M., Schönberger, H. J. M., & Spreen, M. (2007). *Onttrekkingen en recidives tijdens verlof gedurende detbs-behandeling nader bekeken.* Utrecht: Expertisecentrum Forensische Psychiatrie (EFP).

Hillbrand, M. (1995). Aggression against self and aggression against others in violent psychiatric patients. *Journal of Consulting and Clinical Psychology, 63*(4), 668-671.

Hogan, N., & Ennis, L. (2010). Assessing risk for forensic psychiatric inpatient violence: A meta-analysis. *Open Access Journal of Forensic Psychology, 2*, 137-147.

Hoptman, M. J., Yates, K. F., Patalinjug, M. B., Wack, R. C., & Convit, A. (1999). Clinical prediction of assaultive behavior among male psychiatric patients at a maximum-security forensic facility. *Psychiatric Services, 50*(11), 1461-1466.

Jeandarme, I., Pouls, C., Laender, J. De, Oei, T. I., & Bogaerts, S. (2016). *Field validity of the HCR-20 in forensic medium security units in Flanders*. Manuscript in review.

Kelly, E. L., Subica, A. M., Fulginiti, A., Brekke, J. S., & Novaco, R. W. (2015). A cross-sectional survey of factors related to inpatient assault of staff in a forensic psychiatric hospital. *Journal of Advanced Nursing, 71*(5), 1110-1122. doi: 10.1111/jan.12609

Klerx-Van Mierlo, F., & Bogaerts, S. (2011). Vulnerability factors in the explanation of workplace aggression: The construction of a theoretical framework. *Journal of Forensic Psychology Practice, 11*(4), 265-292. doi: 10.1080/15228932.2011.562826

Kraan, J. van der, Verkes, R. J., Goethals, K., Vissers, A., Brazil, I., & Bulten, E. (2014). Substance use disorders in forensic psychiatric patients. *International Journal of Law and Psychiatry, 37*(2), 155-159. doi: 10.1016/j.ijlp.2013.11.009

Kunst, M., Bogaerts, S., & Winkel, F. (2009). Peer and inmate aggression, type D-personality and post-traumatic stress among Dutch prison workers. *Stress and Health, 25*(5), 387–395. doi: 10.1002/smi.1247

Lammers, S. M. M., Soe-Agnie, S., Haan, H. A. de, Bakkum, G. A. M., Pomp, E. R., & I., N. H. L. (2014). Middelengebruik en criminaliteit: Een overzicht [Substance use and criminality: A review]. *Tijdschrift voor Psychiatrie, 56*(1), 32-39.

Linhorst, D. M., & Scott, L. P. (2004). Assaultive behavior in state psychiatric hospitals: Differences between forensic and nonforensic patients. *Journal of Interpersonal Violence, 19*(8), 857-874. doi: 10.1177/0886260504266883

Long, C., Dolley, O., & Hollin, C. (2012). Engagement in psychosocial treatment: Its relationship to outcome and care pathway progress for women in medium-secure settings. *Criminal Behaviour and Mental Health, 22*(5), 336-349. doi: 10.1002/cbm.1824

McDermott, B. E., Dualan, I. V., & Scott, C. L. (2011). The predictive ability of the Classification of Violence Risk (COVR) in a forensic psychiatric hospital. *Psychiatric Services, 62*(4), 430-433. doi: 10.1176/appi.ps.62.4.430

McDermott, B. E., Edens, J. F., Quanbeck, C. D., Busse, D., & Scott, C. L. (2008). Examining the role of static and dynamic risk factors in the prediction of inpatient violence: Variable- and person-focused analyses. *Law and Human Behavior, 32*(4), 325-338. doi: 10.1007/s10979-007-9094-8

McDermott, B. E., Quanbeck, C. D., Busse, D., Yastro, K., & Scott, C. L. (2008). The accuracy of risk assessment instruments in the prediction of impulsive versus predatory aggression. *Behavioral Sciences & the Law, 26*(6), 759–777. doi: 10.1002/bsl.842

McMurran, M., & Theodosi, E. (2007). Is treatment non-completion associated with increased reconviction over no treatment? *Psychology, Crime & Law, 13*(4), 333-343. doi: 10.1080/10683160601060374

Mudde, N., Nijman, H., Hulst, W. van der, & Bout, J. van den (2011). Het voorspellen van agressie tijdens de behandeling van forensisch psychiatrische patiënten aan de hand van de HCR-20 [Predicting aggression during the treatment of forensic psychiatric patients by means of the HCR-20]. *Tijdschrift voor Psychiatrie, 53*(10), 705-713.

Neal, T. M. S., Miller, S. L., & Shealy, R. C. (2015). A field study of a comprehensive violence risk assessment battery. *Criminal Justice and Behavior, 42*(9). doi: 10.1177/0093854815572252

Nicholls, T. L., Brink, J., Greaves, C., Lussier, P., & Verdun-Jones, S. (2009). Forensic psychiatric inpatients and aggression: An exploration of incidence, prevalence, severity, and interventions by gender. *International Journal of Law and Psychiatry, 32*(1), 23-30. doi: 10.1016/j.ijlp.2008.11.007

Nijman, H. L. I., Muris, P., Merckelbach, H. L. G. J., Palmstierna, T., Wistedt, B., Vos, A. M., . . . Allertz, W. (1999). The Staff Observation Aggression Scale-Revised (SOAS-R). *Aggressive Behavior, 25*(3), 197–209.

O'Reilly, K., Donohoe, G., Coyle, C., O'Sullivan, D., Rowe, A., Losty, M., . . . Kennedy, H. G. (2015). Prospective cohort study of the relationship between neuro-cognition, social cognition and violence in forensic patients with schizophrenia and schizoaffective disorder. *BMC Psychiatry, 15*, 155. doi: 10.1186/s12888-015-0548-0

Ogloff, J. R., & Daffern, M. (2006). The dynamic appraisal of situational aggression: An instrument to assess risk for imminent aggression in psychiatric inpatients. *Behavioral Sciences & the Law, 24*(6), 799-813. doi: 10.1002/bsl.741

Palmstierna, T., & Wistedt, B. (1987). Staff observation aggression scale, SOAS: Presentation and evaluation. *Acta Psychiatrica Scandinavica, 76*(6), 657-663.

Pedersen, L., Ramussen, K., & Elsass, P. (2012). HCR-20 violence risk assessments as a guide for treating and managing violence risk in a forensic psychiatric setting. *Psychology, Crime & Law, 18*(8), 733-743. doi: 10.1080/1068316X.2010.548814

Pouls, C., Jeandarme, I., & Habets, P. (2014). Risicotaxatie bij daders met een verstandelijke beperking. Eerste toepassing Nederlandstalige VRAG [Risk assessment in offenders with intellectual disabilities: A first application of the Dutch VRAG]. *De Psycholoog, 1*, 42-51.

Quanbeck, C. (2006). Forensic psychiatric aspects of inpatient violence. *The Psychiatric Clinics of North America, 29*(3), 743-760. doi: 10.1016/j.psc.2006.04.011

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk.* Washington, DC.: American Psychological Association.

Rasmussen, K., & Levander, S. (1996). Individual rather than situational characteristics predict violence in a maximum security hospital. *Journal of Interpersonal Violence, 11*(3), 376-390.

Rogers, P., Watt, A., Gray, N., MacCulloch, M., & Gournay, K. (2002). Content of command hallucinations predicts selfharm but not violence in a medium secure unit. *The Journal of Forensic Psychiatry, 2*, 251-262. doi: 10.1080/09585180210150096

Schuringa, E., Spreen, M., & Bogaerts, S. (2014). Inter-rater and test-retest reliability, internal consistency, and factorial structure of the Instrument for Forensic Treatment Evaluation. *Journal of Forensic Psychology Practice, 14*(2), 127-144. doi: 10.1080/15228932.2014.897536

Sedgwick, O., Young, S., Das, M., & Kumari, V. (2016). Objective predictors of outcome in forensic mental health services-a systematic review. *CNS Spectrums*, 1-15. doi: 10.1017/S1092852915000723

Snowden, R. J., Gray, N. S., Taylor, J., & Fitzgerald, S. (2009). Assessing risk of future violence among forensic psychiatric inpatients with the Classification of Violence Risk (COVR). *Psychiatric Services, 60*(11), 1522-1526. doi: 10.1176/appi.ps.60.11.1522

Spreen, M., Brand, E., Horst, P. Ter, & Bogaerts, S. (2014). *Handleiding en methodologische verantwoording HKT-R, Historisch, Klinische en Toekomstige-Revisie*. Groningen: Dr. S. van Mesdag.

Thomas, S. D., Daffern, M., Martin, T., Ogloff, J. R., Thomson, L. D., & Ferguson, M. (2009). Factors associated with seclusion in a statewide forensic psychiatric service in Australia over a 2-year period. *International Journal of Mental Health Nursing, 18*(1), 2-9. doi: 10.1111/j.1447-0349.2008.00567.x

Vitacco, M. J., Rybroek, G. J. Van, Rogstad, J. E., Yahr, L. E., Tomony, J. D., & Saewert, E. (2009). Predicting short-term institutional aggression in forensic patients: A multi-trait method for understanding subtypes of aggression. *Law and Human Behavior, 33*(4), 308-319. doi: 10.1007/s10979-008-9155-7

Vojt, G., Thomson, L. D. G., & Marshall, L. A. (2013). The predictive validity of the HCR-20 following clinical implementation: Does it work in practice? *The Journal of Forensic Psychiatry & Psychology, 24*(3), 371-385. doi: 10.1080/14789949.2013.800894

Walters, G. D. (2003). Predicting institutional adjustment and recidivism with the Psychopathy Checklist factor scores: A meta-analysis. *Law and Human Behavior, 27*(5), 541-558.

Wang, E. W., & Diamond, P. M. (1999). Empirically identifying factors related to violence risk in corrections. *Behavioral Sciences & the Law, 17*(3), 377-389.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence* (Version 2). Burnaby, BC: Simon Fraser University.

Whittington, R., & Richter, D. (2006). From the individual to the interpersonal: environment and interaction in the escalation of violence in mental health settings. In D. Richter & R. Whittington (Eds.), *Violence in mental health settings: Causes, consequences, management* (pp. 47-68). New York, NY: Springer.

Wilson, C. M., Desmarais, S. L., Nicholls, T. L., Hart, S. D., & Brink, J. (2013). Predictive validity of dynamic factors: Assessing violence risk in forensic psychiatric inpatients. *Law and Human Behavior, 37*(6), 377-388. doi: 10.1037/lhb0000025

Woods, P., & Ashley, C. (2007). Violence and aggression: A literature review. *Journal of Psychiatric and Mental Health Nursing, 14*(7), 652-660. doi: 10.1111/j.1365-2850.2007.01149.x

Woods, P., Olver, M., & Muller, M. (2015). Implementing violence and incident reporting measures on a forensic mental health unit. *Psychology, Crime & Law, 21*(10), 973-986. doi: 10.1080/1068316X.2015.1077247

Forensic psychiatric patients with comorbid psychopathy: Double trouble?

Accepted pending minor revisions

7

Jeandarme, I., Pouls, C., Oei, T. I., & Bogaerts, S.

ABSTRACT

Clinicians are reluctant to treat patients with psychopathy despite their need for intensive care and supervision, not only because of the (supposedly) limited chances of success but also because therapy-interfering behavior is expected. The current study focused on inpatient disruptive behavior in mentally disordered offenders who were being treated in a medium security unit. Patients (*N* = 224) were assessed using the Psychopathy Checklist-Revised total, factor and facet scores and divided into three groups depending on the presence of low, medium and high psychopathy traits. Associations between psychopathy and criminogenic risk and need factors were analyzed. Additionally, the association between psychopathy and therapy-interfering behavior (non-compliance, drop-out, institutional misconduct) was investigated with correlational and logistic regression analyses. The results showed that psychopathy was associated with greater risk, needs and therapy-interfering behavior. Factor 2 predicted institutional misconduct, whereas Factor 1 predicted drop-out from treatment. The study highlights the importance of responsive treatment climates in retaining this difficult-to-treat group in treatment.

INTRODUCTION

Under Belgian law, *internment* is a safety measure for offenders who are found not guilty for reason of insanity (NGRI). As in other countries (Every-Palmer et al., 2014; Salize & Dressing, 2007), this specific legislation allows mentally disordered offenders (MDOs) to be transferred to (forensic) psychiatric facilities for mandatory treatment. In line with research on non-mentally ill offenders, antisocial and/or psychopathic personality traits are also strong risk factors for recidivism in MDOs (Bonta, Blais, & Wilson, 2014). Comorbid psychopathic traits in MDOs are associated with a more impulsive and coercive and less compliant interpersonal styles (Fullam & Dolan, 2006), premeditated aggression (Bo et al., 2013), and poor criminal outcomes (Newhill, Vaughn, & DeLisi, 2010; Richards, Casey, & Lucente, 2003; Tengström, Hodgins, Grann, Langström, & Kullgren, 2004).

Following the Risk-Need-Responsivity (RNR) principles (Andrews & Bonta, 2010), forensic psychiatric treatment should focus intensively on offenders with high psychopathic traits – typically measured with the Psychopathy Checklist-Revised (PCL-R) (Hare, 2003) – and target dynamic criminogenic risk factors, such as impulsivity, hostility, poor regulation of aggressive behavior, lack of empathy and lack of realistic future goals (Bogaerts, Polak, Spreen, & Zwets, 2012). In addition, treatment should address responsivity considerations, such as superficial charm, narcissism, sensation seeking and emotional defects (Andrews & Bonta, 2010). However, addressing these concerns is easier said than done. Clinicians are often very reluctant to accept MDOs with high psychopathy traits for treatment. As noted by Salekin, Worley, and Grimes (2010), regardless of how limited the data may be (D'Silva, Duggan, & McCarthy, 2004; Felthous, 2015; Reidy, Kearns, & DeGue, 2013), clinicians continue to maintain strong opinions about the treatability of psychopathy, ranging from the belief that treatment will have no effect, that it will worsen the condition and to that it could potentially better the outcome.

Furthermore, it is generally believed that high levels of psychopathy guarantee a problematic course of treatment and thus present a serious challenge for therapists. Skeem, Manchak, and Peterson (2011) compared these patients with poor students in the classroom: verbally combative, hostile, prone to break rules and not motivated to cooperate with treatment. According to Wong, Gordon, and Gu (2007), one of the most daunting responsivity factors in

treatment is treating unmotivated, non-adherent and treatment-resistant clients such as many patients with psychopathy or (antisocial) personality disorder.

Therapy-interfering behavior while treating MDOs with psychopathic traits

Meta-analytic research findings in different settings – including forensic psychiatric settings – have indicated that the PCL-R total score is associated with broadly defined institutional misconduct (r_w = .29) and to a lesser extent, with physical violence (r_w = .17) (Guy, Edens, Anthony, & Douglas, 2005). PCL-R Factor 2 was moderately predictive of institutional adjustment, whereas Factor 1 showed a less robust association (Walters, 2003). Of the PCL-R's four facet scores, Facet 4 was the strongest and most incrementally valid predictor of institutional aggression in forensic psychiatric settings (Walters & Heilbrun, 2010). Heilbrun et al. (1998) noted that a relationship between psychopathy and inpatient violence was present during the first two months of forensic psychiatric hospitalization but then disappeared during the last two months, suggesting a need to challenge authority in new situations. However, Rice, Harris, and Cormier (1992) found that psychopaths showed poorer adjustments (more seclusions due to violent or disruptive behavior and more engagement in any disruptive or counter-therapeutic behavior), both in their first *and* in their last year of treatment. Gacono, Meloy, Speth, and Roske (1997) found a relationship between the PCL-R total and factor scores and escapes from a high-security forensic hospital, whereas Cullen et al. (2015) found no association in a forensic psychiatric setting, and psychopathy was not found to predict attempts to escape in schizophrenic patients in particular (Dolan & Davies, 2006). Patients with psychopathy were also more likely to file a complaint against their treatment and treatment providers, which may be yet another form of acting out (Dolan & Millington, 2002).

Additionally, patients with psychopathy tend to devote less effort to their treatment programs (Ogloff, Wong, & Greenwood, 1990), unless it concerns activities that are considered more "fun" such as sports (Hildebrand & de Ruiter, 2012). Patients with psychopathy are non-compliant, behave more negatively and show less improvement compared to non-psychopathic offenders (Hill, Rogers, & Bickford, 1996; Reiss, Grubin, & Meux, 1999). These treatment characteristics have been primarily associated with PCL-R Factor 1 (Hobson, Shine, & Roberts,

2000; Morrissey, Mooney, Hogue, Lindsay, & Taylor, 2007; Olver, Lewis, & Wong, 2013), with the exception of one study (Hildebrand & de Ruiter, 2012).

Given the above-mentioned problems with non-compliance and institutional misconduct and the resulting security concerns, the increased treatment drop-out rates may not be surprising. A meta-analysis performed by Olver, Stockdale, and Wormith (2011) found that having an antisocial personality disorder or psychopathy (both the diagnosis and dimensional PCL-R scores) predicted attrition. In male sex offenders, PCL-R Facet 2 showed significant unique contributions to drop-out, whereas none of the other facets did (Olver & Wong, 2011). In female patients with substance abuse, a shorter treatment stay was associated with PCL-R Factor 1 personality characteristics (Richards et al., 2003).

In summary, the studies cited above indicate that patients with psychopathy have a difficult treatment process that places a significant burden on treatment settings. In addition to the noted common areas of concern, such as poor motivation to change and institutional misconduct, manipulation and deceit or the risk of being fooled by psychopathic patients have also been noted (Salekin et al., 2010).

Current study

The objective of this study was to investigate the relations between psychopathy measured by the PCL-R and indicators of therapy-interfering behavior, namely institutional misconduct, non-compliance and drop-out from treatment. We controlled for other risk/need factors because these may contribute to the associations of interest. More specifically, we aimed to test the following hypotheses. First, we anticipated that the PCL-R total scores would be associated with institutional misconduct, non-compliance and drop-out. Second, we hypothesized that Factor 2 and Facet 4 would be associated with institutional misconduct and Factor 1 and Facet 2 with drop-out. Third, we assumed that although psychopathy would be associated with greater risks and needs, it would still independently increase the risk of therapy-interfering behavior after controlling for other needs and risk factors.

MATERIALS AND METHODS

Setting

The current study is part of a larger study that collected data from 531 patients who were admitted between 2001 and 2010 to one of three medium security units (MSUs) in Flanders (the University Psychiatric Center Sint-Kamillus at Bierbeek, the Public Psychiatric Care Center at Rekem and the Psychiatric Center St. Jan Baptist at Zelzate). Only patients who were assessed with the PCL-R (*N* = 224) were included. To be admitted, an offender must have been found not guilty for reason of insanity and been subjected to the internment measure, which is a safety measure used to protect society from further offending. Other inclusion criteria for MSU admission included being an adult, speaking Dutch, having some degree of motivation and ability to learn, and having an average security and risk level (Moens & Pauwelyn, 2012). Primary psychopathy was one of the exclusion criteria for MSU admission, whereas comorbid psychopathic traits were addressed as responsivity factors. Two out of the three institutions organized their wards based on the patient's key psychopathological disorder (psychotic and personality disorder), whereas the third institution mainly structured the wards according to the patient's progress in treatment. Only one institution included female MDOs. All units used cognitive behavioral therapeutic approaches within a relapse prevention framework. The overall treatment aim was primarily to reduce the patients' risk level in order to transfer them to a lower risk and security level. Following the RNR principles, treatment focused on criminogenic needs (Bonta et al., 2014). Medium security treatment was intensive and lengthy, with an average treatment duration of about two years. The treatment was not specifically designed to treat psychopathy, which was addressed as a responsivity factor.

Participants

The average age of the patients at the time of MSU admission was 34 years (*SD* = 8.76, range = 19.3–60.7), while the average age at the time the PCL-R was administered was 35 years (*SD* = 8.93, range = 18.6–59.8). The average treatment duration was 807 days (*SD* = 519 days, range = 39–2729 days). Patients were predominantly male (98.7%, *n* = 221) and of Belgian nationality (91.1%, *n* = 204). The internment measure was imposed in response to (sexual) violent offenses (79.5%, *n*

= 178), property crimes (16.5%, *n* = 37), drug offenses (1.3%, *n* = 3), hands-off sex offenses (1.3%, *n* = 3), and other offenses (1.3%, *n* = 3). The average age at first conviction or internment measure was 23 years (*SD* = 7.02, range = 10.2–53.6). On average, the sample had 6.9 (*SD* = 5.96, range = 1–38) previous sentences for general offenses and 2.5 (*SD* = 1.92, range = 0–8) previous sentences for violent crimes.

Cumulative diagnoses based on the Diagnostic and Statistical Manual of mental disorders-IV(-Text Revision) (DSM-IV(-TR); American Psychiatric Association, 2000) were determined by the first author in consultation with the attending psychiatrists on the basis of all available file information. More than two-thirds (78.1%, *n* = 175) had a personality disorder, almost half of which was antisocial personality disorder (42.9%, *n* = 75). On Axis I, the most common diagnoses were substance misuse (58.9%, *n* = 132) and psychotic disorders (37.5%, *n* = 84). The average IQ score (*n* = 153) based on the Wechsler Adult Intelligence Scale-III (WAIS-III; Wechsler, 2000) was 81.8 (*SD* = 15.83, range = 48–138). The mean Historical Clinical Risk Management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997) score (*n* = 135) was 25.3 (*SD* = 5.17, range = 11–36).

Instruments

The PCL-R (Hare, 2003; Dutch translation: Vertommen, Verheul, de Ruiter, & Hildebrand, 2002) is the most common and best established method for assessing psychopathy, including in MDO populations (Hart & Hare, 1989). Ideally, the PCL–R is scored using file information and a semi-structured interview with the offender. The tool includes 20 items that are scored on a 3-point scale, where 0 = not present, 1 = perhaps or to some extent present, and 2 = definitely present. Total scores vary between 0 and 40, with higher scores indicating more psychopathic traits. Categorically, a cut-off score of 30 was determined by the developer of the instrument (Hare, 2003). In Europe, a score of 25 or more is considered indicative of psychopathy (Cooke & Michie, 1999). Several studies have supported the dimensional use of the PCL-R score instead of a categorical diagnosis (Edens, Marcus, Lilienfeld, & Poythress, 2006; Marcus, Lilienfeld, Edens, & Poythress, 2006). The initial two-factor model (Factor 1: manipulative and callous unemotional traits and Factor 2: impulsive and antisocial lifestyle) was later subdivided into four facets: Factor 1 was divided into an interpersonal (Facet 1) and an affective facet (Facet 2), and Factor 2 was

divided into a parasitic lifestyle facet (Facet 3) and a criminal history facet (Facet 4). In this model, the formerly separate item *Criminal versatility* was included under Facet 4 (Cooke & Michie, 1999).

In a previous study using a partly overlapping sample, large individual differences were found in addition to moderate interrater reliability (IRR), i.e., $ICC_{A,1}$ = .42 for the PCL-R total score (Jeandarme, Edens, Habets, Bruckers, Oei, & Bogaerts, in press).

Procedure

The PCL-R scores in the current study were field scores, obtained as part of routine clinical practice. The referral indicators for performing a PCL-R evaluation included mainly treatment amenability, lengthy criminal history, antisocial personality disorder, history of predatory violence and potential for violence. When multiple scores were available, the first administered PCL-R was included. All evaluators who scored the PCL-R had a master's or medical degree and were trained by experienced trainers in a manner consistent with the procedures outlined in the test manual (Hare, 2003). Unfortunately, the scoring procedure (whether conducted with or without an interview) was not noted in the files. However, it is likely that the PCL-R was scored with an interview in most cases because this is the method most commonly applied in the field, compared to research studies, which tend to solely rely on file information (Hawes, Boccaccini, & Murrie, 2013). Based on an informal survey of hospital and prison staff conducted by the authors, an interview was typically performed unless the patient refused to participate.

Criminal history variables were extracted from the Central Criminal Record. Incarceration periods were calculated based on extraction from the administrative prison registration system. The HCR-20 scores in the current study were field validity scores. A recent study using a partly overlapping population showed mixed rater consistencies, with IRR coefficients between the field validity scores and research scores ranging from .58 to .84 (Jeandarme, Pouls, De Laender, Oei, & Bogaerts, in press). Other population data were collected from treatment records and judicial files from the Commissions for the Protection of Society (CPSs). These commissions are dedicated bodies chaired by a judge that function as a court of law. The CPSs are responsible for the execution of internment measures and are kept up-to-date on the course of treatment. The patients selected for the current study were on conditional release that was contingent on a number of strict general conditions, such as an obligation to engage in inpatient treatment and

an obligation to stop any criminal activities, as well as specific individual conditions, such as being prohibited from having contact with the victim. Incidents occurring during treatment were reported to the CPS on a regular basis. These judicial incident reports were used in the current study to assess therapy-interfering behavior.

The study was approved by the Ethics committee of Antwerp University Hospital on January 24, 2011.

Outcome measures

The following variables were considered indicators of therapy-interfering behavior (TIB) during inpatient treatment: 1) treatment drop-out, 2) non-compliance and 3) institutional misconduct. Non-compliance was defined as a report to the CPS that the treatment rules had not been respected (such as the use of alcohol during treatment or failure to cooperate with treatment). Institutional misconduct was defined as 1) absconding (such as escaping from the institution or absconding from supervised or unsupervised leave), 2) violating individually formulated judicial conditions, and 3) engaging in offense-related behavior, which was defined as incidents coded under offending categories in the Belgian penal code, regardless of whether they led to further prosecution or sentencing. This definition of the outcome measure was selected because it was highly likely that a public prosecutor in the Belgian internment system would forego prosecution, since the patients were already interned for an indefinite period and residing within an MSU. As a result, using only the official recidivism data would have resulted in an artificially low base rate. These offense-related incident reports were divided into five categories: violent (ranging from threats to murder or rape), non-violent sexual, drug-related (such as drug use or trafficking), property (such as theft), and other offenses (such as destruction of property or traffic violations).

Analyses

The data analysis was performed in SPSS version 22. Three groups with different levels of psychopathy (low = ≤ 15, medium = 16–24, high = ≥ 25) were constructed. The differences between low, medium, and high psychopathic patients were evaluated using Chi-square or Fischer's Exact test for categorical variables and Cramer's *V* as a measure of association strength. One-way ANOVA (for normally distributed data) and Mann-Whitney U tests (for non-normally

distributed data) were used for continuous variables, and *r* effect sizes were calculated. The relationship between PCL-R total, factor and facet scores and the normally distributed continuous variables were examined using Pearson *r* correlations and point-biserial correlations for dichotomous variables. We used Cohen's (1988) conventions to interpret the effect sizes, with correlations of .10 indicating a weak association, .30 indicating a moderate association, and .50 indicating a strong or large association.

Logistic regression analyses were used to determine whether 1) the PCL-R total score was related to TIB, 2) the PCL-R Factor 1 and/or Factor 2 were associated with TIB and 3) any of the four PCL-R facets were associated with TIB.

Missing data were addressed using pairwise deletion. There were missing data for IQ score and for HCR-20 scores, as noted in the participants section. Additionally, PCL-R item scores were not always available. Therefore, the facet analyses were based on smaller samples, as noted in Table 1.

RESULTS

Descriptive statistics are displayed in Table 1. Almost 10% of the patients (8.9%, *n* = 20) had a score of 30 or more, and one-third of the patients (33.5%, *n* = 75) had a score of 25 or more. The correlations between factors and facet scores are also shown in Table 1.

Table 1. Total, factor and facet scores of the PCL-R, and corresponding correlations

	N	*M*	*SD*	Range	Total	Factor 1	Factor 2	Facet 1	Facet 2	Facet 3	Facet 4
Total	224	21.1	6.58	6–36.8	-						
Factor 1	215	8.6	3.40	1.1–16	.73***	-					
Factor 2	213	10.3	3.86	1–18	.79***	.23**	-				
Facet 1	173	3.1	2.14	0–8	.62***	.86***	.16*	-			
Facet 2	174	5.6	1.95	0–8	.59***	.82***	.20**	.41***	-		
Facet 3	171	6.0	2.57	0–10	.73***	.25**	.89***	.20**	.22**	-	
Facet 4	164	5.5	2.59	0–10	.67***	.12	.81***	.08	.13	.48***	-

*p < .05. ** p < .01. *** p < .001.

Is psychopathy associated with greater criminogenic risk/needs?

Table 2 shows the comparison of patients with low (*n* = 54), medium (*n* = 95), and high (*n* = 75) PCL-R total scores with respect to criminal history, risk assessment scores and clinical diagnoses. With the exception of the C-scale of the HCR-20 and the clinical diagnosis of substance use disorder, all other risk factors differed significantly between the three levels of psychopathy, with higher levels of psychopathy showing associations with more risk factors.

Table 2. Comparison of patients with low, medium, and high total PCL-R scores regarding criminal history, risk assessment scores, and clinical diagnoses

	PCL-R ≤ 15 (*n* = 54)		PCL 16–24 (*n* = 95)		PCL ≥ 25 (*n* = 75)			
	M	*SD*	*M*	*SD*	*M*	*SD*	*p*	*r*
Criminal history								
Number of sentences	4.6	3.60	6.5	5.4	9.1	7.2	.00***	0.30
Number of sentences for violence	1.9	1.71	2.4	1.78	3.0	2.10	.00**	0.23
Age first sentence	25.9	7.36	23.1	7.42	21.6	5.67	.00**	0.23
Duration incarceration	1188.6	1108.25	1593.4	1203.00	2042.8	1919.03	.00**	0.22
Risk assessment scores								
HCR-20 C-scale	4.7	2.23	4.8	1.83	4.8	1.48	.95	
HCR-20 R-scale	5.1	2.18	6.5	2.14	5.9	1.87	.00**	0.27
	%	*n*	%	*n*	%	*n*	*p*	Cramer's V
Clinical diagnosis								
Personality disorder	57.4	31	87.4	83	81.3	61	.00***	.29
Substance misuse	48.1	26	63.2	60	61.3	46	.18	

Note. PCL-R = Psychopathy Checklist-Revised; HCR-20 = Historical, Clinical, Risk management-20.
*$p < .05$. ** $p < .01$. *** $p < .001$.

Next, Pearson and point-biserial correlations were computed for the PCL-R total, factor and facet scale scores regarding offense-related variables (age at first sentence, number of sentences, number of sentences for a violent offence, duration of incarceration periods), clinical variables (presence of a personality disorder and presence of substance use disorder), and scores on the HCR-20 C-scale and R-scale. As shown in Table 3, the PCL-R total score correlated positively with the number of sentences, $r(222) = .27$, $p < .001$, number of sentences for a violent crime, $r(222) = .21$, $p = .002$ and the duration of incarceration, $r(222) = .27$, $p < .001$ and showed negative correlations with age at first sentence, $r(222) = -.29$, $p < .001$. In addition, the PCL-R total score

correlated positively with having a personality disorder, $r(222) = .23$, $p < .001$, and the HCR-20 R-scale score, $r(134) = .19$, $p = .03$. With the exception of substance misuse and the HCR-20 C-scale, the PCL-R total score was associated with all other criminogenic factors, with correlations in the small range. Few (small) associations were found for Factor 1, whereas Factor 2 was consistently associated with all variables (except for the HCR-20 C-scale) in the small to medium range. The largest associations were found for Facet 4, with half of the correlations within the medium range.

Table 3. Correlations of PCL-R total, factor and facet scores, and criminogenic risk factors

	PCL-R Total	PCL-R Factor 1	PCL-R Factor 2	PCL-R Facet 1	PCL-R Facet 2	PCL-R Facet 3	PCL-R Facet 4
Number of sentences	.27**	.07	.26**	.08	.05	.29**	.32**
N	224	215	213	172	173	170	163
Number of sentences for violence	.21**	.06	.17*	.04	.13	.14	.33**
N	224	215	213	172	173	170	163
Age at first sentence	-.29**	-.02	-.40**	-.01	.04	-.32**	-.47**
N	224	215	213	172	173	170	163
Duration incarceration	.27**	.16*	.24**	.14	.14	.14	.35**
N	224	215	213	172	173	170	163
Personality disorder: yes/no	.23**	.17*	.19**	.23**	.05	.17*	.14
N	224	215	213	172	173	170	163
Substance misuse disorder: yes/no	.08	-.05	.16*	-.02	-.12	.07	.21**
N	224	215	213	172	173	170	163
HCR C-scale	.08	.14	.01	.02	.26**	.07	-.04
N	138	135	133	116	116	115	109
HCR R-scale	.19*	.16	.18*	.08	.13	.28**	.09
N	136	133	131	114	114	113	107

Note. PCL-R = Psychopathy Checklist-Revised.
*$p < .05$. ** $p < .01$. *** $p < .001$.

Is psychopathy associated with therapy-interfering behavior?

In Table 4, details are provided concerning the TIB of patients with low, medium, and high levels of psychopathic traits. As seen from this table, there was an increase in practically all TIBs as the level of psychopathy increased, with significant differences for the three broad TIB categories (i.e., drop-out, non-compliance and institutional misconduct). More specifically, only one

subcategory of institutional misconduct, absconding, was significantly related to the level of psychopathy, whereas none of the other subcategories showed a relationship. The increase in TIB was most evident when investigating the number of patients involved in institutional misconduct, which increased from 35.2% (*n* = 19) in patients with low psychopathy traits to 62.1%

Table 4. Comparison among patients with low, medium, and high total PCL-R scores regarding therapy-interfering behavior

		PCL-R ≤ 15 (*n* = 54)		PCL 16–24 (*n* = 95)		PCL ≥ 25 (*n* = 75)			
	N	*n*	%	*n*	%	*n*	%	*p*	Cramers' V
Drop-out	186	8	17.4	31	37.8	33	56.9	.00***	.30
Non-compliance	224	15	27.8	31	32.6	36	48.0	.04*	.17
Institutional misconduct	224	19	35.2	59	62.1	48	64.0	.00**	.24
Absconding	224	9	16.7	35	36.8	30	40.0	.01*	.20
Violation juridical conditions	224	3	5.6	6	6.3	5	6.7	1.00	
Offense-related	224	16	29.6	44	46.3	32	42.7	.13	
Drugs	224	5	9.3	20	21.1	14	18.7	.18	
Property	224	2	3.7	7	7.4	5	6.7	.73	
Sexual hands-off	224	0	0	0	0	1	1.3	.58	
Other	224	2	3.7	9	9.5	6	8.0	.45	
Violence	224	11	20.4	21	22.1	22	29.3	.42	
Physical violence	224	8	14.8	12	12.6	13	17.3	.69	
Verbal violence	224	10	18.5	11	11.6	17	22.7	.15	

Note. PCL-R = Psychopathy Checklist-Revised.
*p < .05. ** p < .01. *** p < .001.

(*n* = 59) and 64.0% (*n* = 48) in patients with medium and high levels of psychopathy, $\chi^2(2) = 12.89$, $p < .001$, Cramer's $V = .24$. Additionally, this finding is striking when investigating the number of patients who dropped out from treatment. As seen in Table 4, these analyses were based on a smaller number of patients because at the census date, 17 patients had not yet been discharged. Of those remaining, 17.4% (*n* = 8) of the patients with low psychopathy failed to complete treatment, whereas more drop-out was observed in the group with medium (37.8%, *n* = 31) and high (56.9%, *n* = 33) levels of psychopathy, $\chi^2(2) = 16.93$, $p < .001$, Cramer's $V = .30$.

In addition to the differences found within the three patient groups, the correlation between the presence of drop-out, non-compliance and institutional misconduct was analyzed for the total, factor and facet scores. Table 5 provides an overview of the correlations and shows that total, factor and facet scores were associated with drop-out. Institutional misconduct was associated with the PCL-R total score ($r(222) = .19$, $p < .01$), Factor 2 ($r(211) = .30$, $p < .001$), Facet 3 ($r(168) = .27$, $p < .001$) and Facet 4 ($r(161) = .29$, $p < .001$), whereas no correlations were found for non-compliance.

Table 5. Correlations of PCL-R total, factor and facet scores, and therapy-interfering behavior

	PCL-R Total	PCL-R Factor 1	PCL-R Factor 2	PCL-R Facet 1	PCL-R Facet 2	PCL-R Facet 3	PCL-R Facet 4
Drop-out	.32***	.25**	.26**	.18*	.20*	.25**	.23**
N	186	177	177	140	140	137	132
Non-compliance	.12	.13	.12	.11	.06	.10	.10
N	224	215	213	172	173	170	163
Misconduct	.19**	.01	.30***	-.05	.07	.27***	.29***
N	224	215	213	172	173	170	163

Note. PCL-R = Psychopathy Checklist-Revised.
*$p < .05$. ** $p < .01$. *** $p < .001$.

Finally, a series of regression analyses were performed to further investigate the independent association of psychopathy with the occurrence of TIB, which was correlated with psychopathy in previous analyses. Scores on the HCR-20 R-scale were not included, because due to the missing data, adding these scores would have seriously limited the sample ($n = 85$). We controlled for the following criminal history variables: number of sentences, number of sentences for a violent crime, age at first sentence, and duration of incarceration periods. Additionally, we controlled for clinical variables, namely personality disorder and substance misuse.

The predictors of drop-out were analyzed using three logistic regressions. In the first regression, the PCL-R total score was used, in the second, the two factors were included, and in the third, the four facets were analyzed. The three regression analyses included 186, 177 and 129 of the 224 patients, respectively. The models correctly classified 69.4%, 70.6% and 66.7% of the

respondents, respectively. Nagelkerke's R^2 values of .24, .24, and .23 for the first, second, and third regressions, respectively, indicated a small to medium relationship between the predictors and drop-out. After controlling for offense-related and clinical variables, the PCL-R total score (β = 0.08, *df* = 1, *p* < .01) and Factor 1 score (β = 0.14, *df* = 1, *p* < .05) remained significantly positively associated with drop-out, whereas none of the facet scores were predictive. Predictors of institutional misconduct were analyzed using the three logistic regressions and included 224, 213 and 159 of the 224 patients, correctly classifying 65.2%, 66.2% and 71.1%, respectively. Nagelkerke's R^2 of .11, .17, and .24, respectively, indicated a small to medium relationship between the predictors and institutional misconduct. After controlling for offense-related and clinical variables, the PCL-R total (β = 0.09, *df* = 1, *p* < .01), Factor 2 (β = 0.16, *df* = 1, *p* < .00) and Facet 4 (β = 0.19, *df* = 1, *p* < .05) scores remained significantly positively associated with misconduct.

Table 6. Logistic regression analyses explaining drop-out and misconduct

								95% CI for Exp(β)	
	Variable	β	S.E.	Wald	*df*	*p*	Exp(β)	Lower	Upper
Drop-out									
Model 1	PCL-R Total	0.09	0.03	9.13	1	.00	1.09	1.03	1.16
	Personality disorder	1.56	0.53	8.64	1	.00	4.77	1.68	13.50
Model 2	PCL-R Factor 1	0.14	0.06	6.53	1	.01	1.15	1.03	1.29
	Personality disorder	1.78	0.80	4.90	1	.03	5.92	1.23	28.53
Model 3	Personality disorder	1.85	0.82	5.08	1	.02	6.36	1.27	31.75
Misconduct									
Model 1	PCL-R Total	0.05	0.24	4.65	1	.03	1.05	1.01	1.10
	Age at first sentence	-0.05	0.02	3.97	1	.05	0.96	0.91	1.00
Model 2	PCL-R Factor 2	0.16	0.05	11.37	1	.00	1.17	1.07	1.28
	Duration of incarceration	0.00	0.00	4.02	1	.05	1.00	1.00	1.00
Model 3	PCL-R Facet 4	0.19	0.09	4.32	1	.04	1.21	1.01	1.46
	Number of sentences for violence	-0.29	0.14	4.20	1	.04	0.75	0.57	0.99

Note. PCL-R = Psychopathy Checklist-Revised.

DISCUSSION

In the current study, the characteristics of MDOs with psychopathic traits were studied regarding TIB during inpatient treatment. The study investigated the PCL-R total score, including its two factors and four facets, in relation to drop-out, institutional misconduct and non-compliance. Psychopathy scores were controlled for related needs and risk factors to ascertain that the associations identified were not confounded by other variables. In summary, psychopathy was associated with greater risk, needs and TIB. Factor 2 predicted institutional misconduct, whereas Factor 1 predicted drop-out from treatment.

First, whether greater criminogenic needs and risk factors differentiated between psychopathic and non-psychopathic MDOs was examined. The results confirmed this hypothesis and were thus in line with similar research with a different population, namely inmates (Simourd & Hoge, 2000). In the current study, patients with psychopathy showed significant involvement with the criminal justice system in criminal history variables (younger age at first sentence, more sentences, more sentences for a violent crime and longer incarceration periods), scored higher on the HCR-20 and showed higher rates of comorbid clinical problems (personality disorder and substance use disorder). Greater risk and need factors were more highly correlated with PCL-R Factor 2 than with Factor 1.

Clinicians are reluctant to admit psychiatric patients with psychopathy because they are skeptical about their chances of successful treatment. Another factor that limits admissions is fear of institutional misconduct and violence; these behaviors negatively influence the treatment process and have an impact on the stability of the facility and can result in high staff turnover (Gow, Choo, Darjee, Gould, & Steele, 2010). Accordingly, whether psychopathy was associated with TIB was investigated. To this end, the three groups with low, medium, and high levels of psychopathy were compared regarding compliance to treatment, institutional misconduct, and drop-out from treatment. Second, correlations between the PCL-R total, factor and facets scores and TIB were analyzed. Finally, significant associations were further explored using regression analyses. The findings partially confirmed that treatment was more difficult for patients with higher levels of psychopathy. When comparing the groups with low, medium, and high psychopathic traits, a small significant difference was found, indicating that patients with more

psychopathic traits were less compliant. However, when investigating the correlations between the PCL-R total, factor and facet scores and non-compliance, no associations were found, which contradicts other research reporting associations with PCL-R total (Hildebrand & de Ruiter, 2012; Hobson et al., 2000; Olver et al., 2013), Factor 1 (Hobson et al., 2000; Olver et al., 2013), and Factor 2 scores (Hildebrand & de Ruiter, 2012). The differences in the operationalization of non-compliance may have contributed to these different findings. Additionally, it should be noted that in the current study, non-compliance might have been underreported.

When investigating the three psychopathy groups with regard to the second type of TIB - institutional misconduct - significant differences were also found, with psychopathic patients showing more misconduct. When further analyzing the different forms of misconduct, the only significant difference was found for absconding, not with offending behavior during treatment. Additionally, as was the case in other single studies (Buffington-Vollum, Edens, Johnson, & Johnson, 2002; Hildebrand, De Ruiter, & Nijman, 2004), psychopathy was not associated with inpatient (physical) violence in the present study. Vitacco et al. (2009) noted that it is important to distinguish between reactive and instrumental violence and found a connection only with the latter. Additionally, the treatment period during which studies are implemented may result in different outcomes. For example, Heilbrun et al. (1998) found a link between psychopathy and violence at the start of treatment, but that association disappeared at a later stage of treatment. The current study, however, did not explore these nuances. The correlational and regression analyses did confirm that the PCL-R total, Factor 2 and Facet 4 scores were associated with the broader category of institutional misconduct, thereby adding to the evidence demonstrating this relationship (Guy et al., 2005; Walters, 2003; Walters & Heilbrun, 2010).

Third, as expected, there was significantly more drop-out in the groups with higher PCL-R scores: more than half of the patients with a PCL-R score equal to or higher than 25 did not finish their treatment. The percentages reported in the literature with respect to drop-out vary from 10 to 90%, with 33% drop-out in forensic psychiatric populations (Olver et al., 2011) and lower drop-out rates in clinical programs (14.7%) compared to ambulatory programs (McMurran & Theodosi, 2007). As we were analyzing a residential treatment program in which drop-out resulted in re-incarceration, the current figures seem to be very high and a reason for concern, since drop-out has been associated with higher recidivism rates (Hanson & Bussiere, 1998; McMurran &

Theodosi, 2007). The correlational analyses showed an association between drop-out and the PCL-R total, both factor and all four facet scores. The regression analyses revealed that only the total and Factor 1 scores were independently related to drop-out. In contrast to a study by Olver and Wong (2011), none of the facets were associated with drop-out.

Considered together, our findings thus confirmed other research stating that treatment should focus on criminogenic PCL-R Factor 2 features while also carefully accounting for PCL-R Factor 1 characteristics to keep patients in treatment (Wong & Olver, 2015).

Study limitations

Some limitations of this study warrant caution when interpreting and generalizing the current findings. First, the PCL-R scores used in the current study were field validity scores with a low IRR. This finding is generally consistent with a growing body of field research that suggests that the high levels of reliability reported in many controlled research studies are not generalizable to practice settings (DeMatteo et al., 2014; Levenson, 2004; Miller, Kimonis, Otto, Kline, & Wasserman, 2012; Neal, Miller, & Shealy, 2015). However, it calls into question how reliable clinical scores truly are. For example, it cannot be ruled out that misconduct during treatment heavily influenced PCL-R scores; i.e., problems of circularity may have occurred. Therefore, prospective studies using clinical and research scores are highly advisable.

Second, TIB was limited to incidents that were reported to the CPS, suggesting that there may have been underreporting especially with regard to non-compliance and verbal violence. The dark number of actual incidents occurring during treatment will likely be higher than the number identified in this study. Additionally, no formal interrater reliability tests were conducted regarding TIB.

Third, it should be noted that only four out of ten patients admitted to the medium security units were assessed with the PCL-R. This may explain the relatively high mean total scores and the high number of patients with a score of 25 or above, especially when compared to the results of another study performed in the French-speaking part of Belgium (*M* = 17.6; Pham, Saloppé, Bongaerts, & Hoebanx, 2007). However, the average PCL-R total score found in the current study was similar to the mean score among forensic psychiatric patients reported by Hare (2003).

Finally, as this study decided to analyze data for the three medium security units together, the impact of possible differences in treatment objectives and climate could not be examined.

Clinical implications

Therapeutic settings involved in the treatment of MDOs with psychopathic traits face major challenges. Psychopathy is linked to institutional maladjustment, lack of motivation, early drop-out from treatment, and slower progression as well as limited treatment outcomes, for example, in terms of reducing recidivism. The behavioral manifestations of these traits can significantly interfere with treatment, as they impede the formation of a good working alliance and therefore must be appropriately managed (Wong et al., 2007). The current study found evidence for the two-component model proposed by Wong, Gordon, Gu, Lewis, and Olver (2012). The Criminogenic component, or Factor 2, was associated with greater criminogenic need/risk factors and institutional misconduct and reflected an established pattern of antisocial behavior and dysfunctional lifestyle both inside and outside the institution. While the Criminogenic component should be the focus of forensic treatment, the Interpersonal component, or Factor 1, is equally important. Factor 1 was associated with drop-out and can thus also contribute to poor outcomes if TIBs are not appropriately managed. As noted by Wong et al. (2012), it is important to distinguish between using risk reduction versus personality change interventions when designing and implementing treatment programs. However, the abovementioned study limitations warrant caution regarding the findings. At this point, no firm conclusions can be drawn from the present study. Further research is needed to corroborate these findings.

Psychopathic features are important, but other individual characteristics can also affect treatment outcomes (Hare, Clark, Grann, & Thornton, 2000). For instance, the present study found that a DSM-IV personality disorder diagnosis also predicted drop-out and institutional misconduct. Moreover, institutional TIB is related not only to patient variables but also to the therapeutic climate and treatment context. Nevertheless, as observed, psychopathic and other personality disorder traits in MDOs may be barriers to forensic psychiatric treatment and can lead to premature interruption or discontinuation of treatment. Clinicians are advised to develop a responsive treatment climate with staff and management who are willing to invest time and effort in these personality disordered patients (for good practices, please see Bulten & Decoene,

2015). Instead of viewing poor motivation as a contraindication to treatment, motivation should be evaluated and innovative treatment studies should be designed to enhance individual's motivation for treatment (Salekin et al., 2010). Previous research has shown that it is important to continue to invest in this population and that establishing a therapeutic alliance, at least for some of these patients, is indeed possible (Polaschek & Ross, 2010). Within-group differences in terms of responsiveness can help determine which psychopathic patients will be most likely to benefit from treatment (Chakhssi, de Ruiter, & Bernstein, 2010).

A meta-analysis revealed that treatment responsivity indicators such as disruptive behavior during treatment and negative treatment attitudes were among the strongest predictors of increased attrition rates, while higher levels of motivation and treatment engagement predicted decreased attrition Olver et al. (2011). Therefore, it is also important to measure progress in therapy in a uniform manner, with tools specifically developed for therapeutic measurement, such as the Instrument for Forensic Treatment Evaluation (IFBE; Schuringa, Spreen, & Bogaerts, 2014).

CONCLUSION

Overall, this study highlighted that even after controlling for criminal history and clinical variables, psychopathy contributed to an increased understanding of TIB. Psychopathy was associated with institutional misconduct and drop-out from treatment, which are considered risk factors for relapse. Although it is understandable that clinicians often prefer pleasant, "YAVIS"-like low-risk patients, the RNR model states that most resources should be deployed to treat more difficult, less compliant patients. Forensic mental health professionals therefore face a great challenge. They must tolerate difficult interpersonal behavior, such as hostility and manipulation, and control their countertransference while still motivating forensic patients who seemingly do not want to change or even stay in treatment. This is clearly not an easy task, but as shown by international research, it is not impossible. Difficult-to-treat should not become synonymous with untreatable.

REFERENCES

Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law, 16*(1), 39-55.

Bo, S., Forth, A., Kongerslev, M., Haahr, U. H., Pedersen, L., & Simonsen, E. (2013). Subtypes of aggression in patients with schizophrenia: The role of personality disorders. *Criminal Behaviour and Mental Health, 23*(2), 124-137. doi: 10.1002/cbm.1858

Bogaerts, S., Polak, M., Spreen, M., & Zwets, A. (2012). High and low aggressive narcissism and anti-social lifestyle in relationship to impulsivity, hostility, and empathy in a group of forensic patients in the Netherlands. *Journal of Forensic Psychology Practice, 12*(2), 147-162. doi: 10.1080/15228932.2012.650144

Bonta, A., Blais, J., & Wilson, H. (2014). A theoretically informed meta-analysis of the risk for general and violent recidivism for mentally disordered offenders. *Aggression and Violent Behavior, 19*, 278–287.

Buffington-Vollum, J., Edens, J. F., Johnson, D. W., & Johnson, J. K. (2002). Psychopathy as a predictor of institutional misbehavior among sex offenders: A prospective replication. *Criminal Justice and Behavior, 29*(5), 497-511. doi: 10.1177/009385402236730

Bulten, E., & Decoene, S. (2015). Management van psychopathie in een klinische context. In W. Canton, L. Claes, L. Gijs, I. Jeandarme, E. Klein-Haneveld, & D. van Beek (Eds.), *Handboek psychopathie en de antisociale-persoonlijkheidsstoornis* (pp. 359-375). Utrecht: De Tijdstroom.

Chakhssi, F., Ruiter, C. de, & Bernstein, D. (2010). Change during forensic treatment in psychopathic versus nonpsychopathic offenders. *The Journal of Forensic Psychiatry & Psychology, 21*(5), 660-682. doi: 10.1080/14789949.2010.483283

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Erlbaum.

Cooke, D. J., & Michie, C. (1999). Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology, 108*(1), 58–68. doi: 10.1037/0021-843X.108.1.58

Cullen, A. E., Jewell, A., Tully, J., Coghlan, S., Dean, K., & Fahy, T. (2015). A prospective cohort study of absconsion incidents in forensic psychiatric settings: Can we identify those at high-risk? *PLoS One, 10*(9), e0138819. doi: 10.1371/journal.pone.0138819

D'Silva, K., Duggan, C., & McCarthy, L. (2004). Does treatment really make psychopaths worse? A review of the evidence. *Journal of Personality Disorders, 18*(2), 163-177.

DeMatteo, D., Edens, J. F., Galloway, M., Cox, J., Smith, S. T., Koller, J. P., & Bersoff, B. (2014). Investigating the role of the Psychopathy Checklist-Revised in United States case law. *Psychology, Public Policy, and Law*, *20*(1), 96-107. doi: 10.1037/a0035452

Dolan, M., & Davies, G. (2006). Psychopathy and institutional outcome in patients with schizophrenia in forensic settings in the UK. *Schizophrenia Research, 81*(2-3), 277-281. doi: 10.1016/j.schres.2005.07.002

Dolan, M., & Millington, J. (2002). The influence of personality traits such as psychopathy on detained patients using the NHS complaints procedure in forensic settings. *Personality and Individual Differences, 33*(6), 955-965.

Edens, J. F., Marcus, D. K., Lilienfeld, S. O., & Poythress, N. G., Jr. (2006). Psychopathic, not psychopath: taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology, 115*(1), 131-144. doi: 10.1037/0021-843X.115.1.131

Every-Palmer, S., Brink, J., Chern, T. P., Choi, W. K., Hern-Yee, J. G., Green, B., . . . Mellsop, G. (2014). Review of psychiatric services to mentally disordered offenders around the Pacific Rim. *Asia-Pacific Psychiatry, 6*(1), 1-17. doi: 10.1111/appy.12109

Felthous, A. R. (2015). The appropriateness of treating psychopathic disorders. *CNS Spectrums, 20*(3), 182-189. doi: 10.1017/s1092852915000243

Fullam, R., & Dolan, M. (2006). The criminal and personality profile of patients with schizophrenia and comorbid psychopathic traits. *Personality and Individual Differences, 40*(8), 1591–1602.

Gacono, C. B., Meloy, J. R., Speth, E., & Roske, A. (1997). Above the law: Escapes from a maximum security forensic hospital and psychopathy. *Journal of the American Academy of Psychiatry and the Law, 25*(4), 547-550.

Gow, R. L., Choo, M., Darjee, R., Gould, S., & Steele, J. (2010). A demographic study of the Orchard Clinic: Scotland's first medium secure unit. *The Journal of Forensic Psychiatry & Psychology, 21*(1), 139-155. doi: 10.1080/14789940903188956

Guy, L. S., Edens, J. F., Anthony, C., & Douglas, K. S. (2005). Does psychopathy predict institutional misconduct among adults? A meta-analytic investigation. *Journal of Consulting and Clinical Psychology, 73*(6), 1056-1064. doi: 10.1037/0022-006X.73.6.1056

Hanson, R. K., & Bussiere, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology, 66*(2), 348-362.

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist*. Toronto, ON: Multi-Health Systems.

Hare, R. D., Clark, D., Grann, M., & Thornton, D. (2000). Psychopathy and the predictive validity of the PCL-R: An international perspective. *Behavioral Sciences & the Law, 18*(5), 623-645.

Hart, S. D., & Hare, R. D. (1989). Discriminant validity of the Psychopathy Checklist in a forensic psychiatric population. *Psychological Assessment, 1*(3), 211-218.

Hawes, S. W., Boccaccini, M. T., & Murrie, D. C. (2013). Psychopathy and the combination of psychopathy and sexual deviance as predictors of sexual recidivism: Meta-analytic findings using the Psychopathy Checklist-Revised. *Psychological Assessment, 25*(1), 233-243. doi: 10.1037/a0030391

Heilbrun, K., Hart, S. D., Hare, R. D., Gustafson, D., Nunez, C., & White, A. J. (1998). Inpatient and postdischarge aggression in mentally disordered offenders: The role of psychopathy. *Journal of Interpersonal Violence, 13*(4), 514-527. doi: 10.1177/088626098013004007

Hildebrand, M., & Ruiter, C. de (2012). Psychopathic traits and change on indicators of dynamic risk factors during inpatient forensic psychiatric treatment. *International Journal of Law and Psychiatry, 35*(4), 276-288. doi: 10.1016/j.ijlp.2012.04.001

Hildebrand, M., Ruiter, C. de, & Nijman, H. (2004). PCL-R psychopathy predicts disruptive behavior among male offenders in a Dutch forensic psychiatric hospital. *Journal of Interpersonal Violence, 19*(1), 13-29. doi: 10.1177/0886260503259047

Hill, C. D., Rogers, R., & Bickford, M. E. (1996). Predicting aggressive and socially disruptive behavior in a maximum security forensic psychiatric hospital. *Journal of Forensic Sciences, 41*(1), 56-59.

Hobson, J., Shine, J., & Roberts, R. (2000). How do psychopaths behave in a prison therapeutic community? *Psychology, Crime & Law, 6*(2), 139-154. doi: 10.1080/10683160008410838

Levenson, J. S. (2004). Reliability of sexually violent predator civil commitment criteria in Florida. *Law and Human Behavior*, *28*(4), 357-368.

Marcus, D. K., Lilienfeld, S. O., Edens, J. F., & Poythress, N. G. (2006). Is antisocial personality disorder continuous or categorical? A taxometric analysis. *Psychological Medicine, 36*(11), 1571-1581. doi: 10.1017/S0033291706008245

McMurran, M., & Theodosi, E. (2007). Is treatment non-completion associated with increased reconviction over no treatment? *Psychology, Crime & Law, 13*(4), 333-343. doi: 10.1080/10683160601060374

Miller, C. S., Kimonis, E. R., Otto, R. K., Kline, S. M., & Wasserman, A. L. (2012). Reliability of risk assessment measures used in sexually violent predator proceedings. *Psychological Assessment*, *24*(4), 944-953. doi: 10.1037/a0028411

Moens, I., & Pauwelyn, L. (2012). *Geen opsluiting, maar sleutels tot re-integratie. Voorstellen voor een gecoördineerd zorgtraject voor geïnterneerden*. Brussel: Zorgnet Vlaanderen.

Morrissey, C., Mooney, P., Hogue, T. E., Lindsay, W. R., & Taylor, J. L. (2007). Predictive validity of the PCL-R for offenders with intellectual disability in a high security hospital: Treatment progress. *Journal of Intellectual and Developmental Disability, 32*(2), 125-133. doi: 10.1080/13668250701383116

Neal, T., Miller, S. L., & Shealy, R. C. (2015). A field study of a comprehensive violence risk assessment battery. *Criminal Justice and Behavior*, *42*(9), 952-968. doi: 10.1177/0093854815572252

Newhill, C., Vaughn, M., & DeLisi, M. (2010). Psychopathy scores reveal heterogeneity among patients with borderline personality disorder. *The Journal of Forensic Psychiatry & Psychology, 21*(2), 202-220. doi: 10.1080/14789940903281157

Ogloff, J. R. P., Wong, S., & Greenwood, A. (1990). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences & the Law, 8*(2), 181-190. doi: 10.1002/bsl.2370080210

Olver, M. E., Lewis, K., & Wong, S. C. (2013). Risk reduction treatment of high-risk psychopathic offenders: The relationship of psychopathy and treatment change to violent recidivism. *Personality Disorders, 4*(2), 160-167. doi: 10.1037/a0029769

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2011). A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. *Journal of Consulting and Clinical Psychology, 79*(1), 6-21. doi: 10.1037/a0022200

Olver, M. E., & Wong, S. (2011). Predictors of sex offender treatment dropout: Psychopathy, sex offender risk, and responsivity implications. *Psychology, Crime & Law, 17*(5), 457-471. doi: 10.1080/10683160903318876

Polaschek, D. L., & Ross, E. C. (2010). Do early therapeutic alliance, motivation, and stages of change predict therapy change for high-risk, psychopathic violent prisoners? *Criminal Behaviour and Mental Health, 20*(2), 100-111. doi: 10.1002/cbm.759

Reidy, D. E., Kearns, M. C., & DeGue, S. (2013). Reducing psychopathic violence: A review of the treatment literature. *Aggression and Violent Behavior, 18*(5), 527-538. doi: 10.1016/j.avb.2013.07.008

Reiss, D., Grubin, D., & Meux, C. (1999). Institutional performance of male 'psychopaths' in a high-security hospital. *The Journal of Forensic Psychiatry, 10*(2), 290-299. doi: 10.1080/09585189908403682

Rice, M. E., Harris, G. T., & Cormier, C. A. (1992). An evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior, 16*(4), 399-412.

Richards, H., Casey, J., & Lucente, S. (2003). Psychopathy And Treatment Response In Incarcerated Female Substance Abusers. *Criminal Justice and Behavior, 30*(2), 251-276.

Salekin, R. T., Worley, C., & Grimes, R. D. (2010). Treatment of psychopathy: A review and brief introduction to the mental model approach for psychopathy. *Behavioral Sciences & the Law, 28*(2), 235-266. doi: 10.1002/bsl.928

Salize, H. J., & Dressing, H. (2007). Admission of mentally disordered offenders to specialized forensic care in fifteen European Union member states. *Social Psychiatry and Psychiatric Epidemiology, 42*(4), 336-342. doi: 10.1007/s00127-007-0159-2

Schuringa, E., Spreen, M., & Bogaerts, S. (2014). Inter-rater and test-retest reliability, internal consistency, and factorial structure of the Instrument for Forensic Treatment Evaluation. *Journal of Forensic Psychology Practice, 14*(2), 127-144. doi: 10.1080/15228932.2014.897536

Simourd, D., & Hoge, R. (2000). Criminal Psychopathy. A Risk-and-Need Perspective. *Criminal Justice and Behavior, 27*(2), 256-272.

Skeem, J. L., Manchak, S., & Peterson, J. K. (2011). Correctional policy for offenders with mental illness: Creating a new paradigm for recidivism reduction. *Law and Human Behavior, 35*(2), 110-126. doi: 10.1007/s10979-010-9223-7

Tengström, A., Hodgins, S., Grann, M., Langström, N., & Kullgren, G. (2004). Schizophrenia and criminal offending. The role of psychopathy and substance use disorders. *Criminal Justice and Behavior, 31*(4), 367-391

Vertommen, H., Verheul, R., Ruiter, C. de, & Hildebrand, M. (2002). *Handleiding bij de herziene versie van Hare's Psychopathie Checklist*. Lisse: Swets Test Publishers.

Walters, G. D. (2003). Predicting institutional adjustment and recidivism with the Psychopathy Checklist factor scores: A meta-analysis. *Law and Human Behavior, 27*(5), 541-558.

Walters, G. D., & Heilbrun, K. (2010). Violence risk assessment and facet 4 of the Psychopathy Checklist: Predicting institutional and community aggression in two forensic samples. *Assessment, 17*(2), 259-268. doi: 10.1177/1073191109356685

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence* (Version 2). Burnaby, BC: Simon Fraser University.

Wong, S. C., Gordon, A., & Gu, D. (2007). Assessment and treatment of violence-prone forensic clients: an integrated approach. *British Journal of Psychiatry Suppl, 49*, s66-74. doi: 10.1192/bjp.190.5.s66

Wong, S. C., Gordon, A., Gu, D., Lewis, K., & Olver, M. (2012). The effectiveness of violence reduction treatment for psychopathic offenders: Empirical evidence and a treatment model. *International Journal of Forensic Mental Health, 11*(4), 336-349. doi: 10.1080/14999013.2012.746760

Wong, S. C., & Olver, M. E. (2015). Risk reduction treatment of psychopathy and applications to mentally disordered offenders. *CNS Spectrums, 20*(3), 303-310. doi: 10.1017/s1092852915000322

Field validity of the HCR-20 in forensic medium security units in Flanders

Psychology, Crime & Law, in press

8

Jeandarme, I., Pouls, C., De Laender, J., Oei, T. I., & Bogaerts, S.

ABSTRACT

Structured risk assessment has become part of routine practice in forensic settings. However, little attention has been paid to the clinical applicability of existing tools. The present research focused on the performance of the Historical Clinical Risk Management-20 (HCR-20) – one of the most commonly used tools for structured professional judgment – in the daily practice of three medium security units in Flanders. Areas under the curve for the prediction of violent recidivism during (*N* = 168) and after (*N* = 105) medium security treatment were non-significant. In addition, analyses showed that the HCR-20 was mainly of interest in identifying low-risk individuals. Further research measuring different aspects of predictive validity in applied settings is recommended.

INTRODUCTION

Empirically based risk assessment using an actuarial or a structured professional judgment (SPJ) approach have become part of routine practice in forensic psychiatry (Fazel, Singh, Doll, & Grann, 2012). In several countries, such as the United Kingdom and the Netherlands, the use of risk assessment is recommended (e.g., Department of Health, 2007; Ministry of Health, 1998) or even mandatory (Dienst Justitiële Inrichtingen, 2014). In Belgium, this is not (yet) the case.

In a meta-review comparing 126 instruments, little variation was found in the predictive validity of risk assessment schemes (Singh & Fazel, 2010). The Historical Clinical Risk Management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997), a SPJ tool, was among the most frequently administered instruments and deemed useful in the development and monitoring of risk management (Hurducas, Singh, de Ruiter, & Petrila, 2014; Singh et al., 2014). An overview of the extensive research with the HCR-20 across different countries, outcome measures (including non-violent recidivism) and populations (including civil psychiatric settings) is provided by Douglas, Shaffer, et al. (2014). Overall, research indicated that the HCR-20 total and scale scores were associated with violence with moderate to moderate/large effect sizes (Douglas & Reeves, 2010). Although only six studies examined the HCR-20 summary risk rating, results supported their utility (Douglas & Reeves, 2010). In addition, O'Shea, Mitchell, Picchioni, and Dickens (2013) found that, compared to all other scales (total, HC15, C5, R5), the summary risk ratings had the largest mean effect size for the prediction of inpatient aggression.

Specifically in forensic psychiatric populations, mostly moderate to large effect sizes were found (e.g., Dolan & Khawaja, 2004; de Vogel & de Ruiter, 2006). However, the review of Hogan and Ennis (2010) showed only small effect sizes for the prediction of inpatient violence. Furthermore, Wilson, Desmarais, Nicholls, Hart, and Brink (2013) found that changes in dynamic risk factors in forensic psychiatric inpatients significantly predicted institutional violence, even after controlling for static risk factors. Regarding inter-rater reliability (IRR), Douglas and Reeves (2010) reported lower IRR coefficients for the summary risk rating (*Mdn* ICC = .66) and Clinical scale (*Mdn* ICC = .74), compared to the Risk Management (*Mdn* ICC = .83) and Historical (Mdn ICC = .83) scale and total score (*Mdn* ICC = .82). Furthermore, Rufino, Boccaccini, and Guy (2011) found an inverse association between subjectivity and rater agreement at item level, with the

items of the C- and R-scales requiring the most subjective judgment. They argue that the relation between scoring subjectivity and predictive validity is not yet clear. Furthermore, a survey revealed that conducting a HCR-20 assessment is time-consuming, with a median time to complete file review, clinical interviews, collection of collateral data, and reporting being 14.2 hours (Green, Carroll, & Brett, 2010). Hurducas et al. (2014) noticed that these and other practical considerations this could lead to questioning the utility of such assessments. However, when assessments are completed valid and reliable, the benefits are numerous including focusing a team on the core issues that directly affect the clinical care and demonstrating gaps in the data relevant to providing insight into risk scenarios (Green et al., 2010).

Most, if not all, of the abovementioned studies were based on assessments by trained researchers in a carefully planned study design. So far, little attention has been paid to the field validity of risk assessment tools in general, and the HCR-20 in specific. A very recent field study using archival data from 230 cases reported "broad null predictive validity findings" for different outcome measures (Neal, Miller, & Shealy, 2015, p. 962). In addition, in a prospective field study, Vojt, Thomson, and Marshall (2013) examined the HCR-20 following the implementation into clinical practice among 109 mentally disordered offenders in Scotland. The HCR-20 consensus score was a poor predictor of – mostly inpatient – violence. One possible explanation for their findings was that the implementation of the HCR-20 led to a reduction of violent behavior due to effective clinical interventions prior to an incident, in turn reducing the predictive accuracy. Evidence supporting this interpretation was found in the low base rate of violence in comparison to a previous study in the same hospital in which the assessors were researchers. A similar conclusion was reached in another implementation study in Denmark where poor to moderate AUCs were found for violent reconvictions and inpatient aggression (Pedersen, Ramussen, & Elsass, 2012). As Vojt et al. (2013) noticed, another explanation of reduced quality and accuracy points to scoring being influenced "by the messy reality of clinical practice and implementation" (p. 383). de Vogel and de Ruiter (2006) partly examined this question and found different risk ratings by clinicians as opposed to researchers. Group leaders conducting the daily supervision and spending most of their time with the patients performed worse in comparison to treatment supervisors and researchers. Overall, the consensus ratings outperformed the individual ratings (de Vogel & de Ruiter, 2006). Furthermore, in the abovementioned studies, clinical teams

attended a training preceding the implementation and organized meetings discussing the ratings, while HCR-20 assessments in 'real life' by field evaluators (i.e., practicing clinicians) will most likely be completed by different raters, with different training backgrounds. Penney, McMaster, and Wilkie (2014) investigated differences in ratings by clinicians and found that psychiatrists and those with formal training in the HCR-20 achieved a higher degree of IRR on the C-subscale, but not on the R-subscale or the summary risk rating as compared to non-psychiatrists and raters without prior training. This difference is likely to increase in actual practice since clinicians have access to different types and amounts of information. It is unclear whether experience in general or experience scoring the HCR-20 is related to scoring accuracy. However, at least one study found a negative correlation between experience and accuracy (Walters, Kroner, & DeMatteo, 2014). As the authors speculated, "experience may breed overconfidence which in turn encourages overestimation of high risk " (p. 298). In addition, contextual pressures such as possible emotional biases may have an effect on scores. For example, group leaders conducting the daily supervision and spending most of their time with the patients tended to give lower HCR-20 scores (de Vogel, de Ruiter, Hildebrand, Bos, & van de Ven, 2004). Individual differences and contextual pressures have been identified as confounding factors in research on other risk assessment schemes as well. For example, there is growing evidence that scoring the Psychopathy Checklist-Revised (PCL-R; Hare, 2003) is affected by the evaluation context, where mean PCL-R scores provided by prosecution-retained experts have been significantly higher compared to experts retained by the defense (DeMatteo et al., 2014). Another drawback in research is that nearly all studies on the predictive validity of risk assessment tools report Receiver Operating Characteristic (ROC), producing the Area Under the Curve (AUC) values, as an outcome measure, which indicates the probability of a randomly selected recidivist having a higher risk classification than a randomly selected non-recidivist. However, this retrospective measure doesn't provide a full picture of the predictive value of a risk assessment tool. Singh (2013) therefore recommends including both components of a tools' predictive validity, namely discrimination and calibration. Discrimination indexes refer to the (retrospective) ability of an instrument to separate those who went on to be violent from those who did not (AUC, sensitivity, specificity), where calibration indexes refer to prospective qualities as to how well a tool's prediction of risk agrees with the actual observed risk (positive predictive value (PPV), negative

predictive value (NPV), number needed to detain (NND), number safely discharged (NSD)). Prospectively orientated statistics such as NND (number of participants judged to be at high risk who need to be detained to prevent a single incident or offense) and NSD (number of participants judged to be at low risk who could be discharged prior to a single incident or offense) might provide more useful information in that they simulate clinical decision-making. However, it should be noted that calibration indices are dependent on base rate figures, with low base rate behaviors leading to over-prediction. In secure units, on average higher base-rates can be expected. In a meta-analysis of Singh, Grann, and Fazel (2011), some of these additional performance indicators were provided for commonly used risk assessment tools. The HCR-20 produced a mean AUC of .70, a PPV of 71% and a NPV of 67%.

Current study

In sum, there is a scarcity of studies examining the use and accuracy of risk assessment tools in daily practice. The aim of the current study was to examine the field validity of the HCR-20 in a forensic psychiatric population.

In accordance with the study of de Vogel and de Ruiter (2004), it was hypothesized that IRR between clinicians and researchers would be lower compared to IRR between researchers.

Based on the (limited) literature investigating the predictive validity of the HCR-20 in daily practice, it was hypothesized that the AUC would most likely be lower compared to the performance in carefully planned research designs (Pedersen et al., 2012; Vojt et al., 2013). As field validity studies using performance indicators other than AUC are currently non-existent, no assumptions could be made regarding sensitivity, specificity, PPV or NPV.

METHOD

Setting

The study was conducted at the three forensic medium security units (MSUs) in Flanders, located in the communities of Bierbeek, Rekem and Zelzate. The establishment of MSUs in 2001 aimed to provide a treatment setting for patients found not guilty for reason of insanity after having committed an offense (NGRI, in Belgium referred to as 'internees'). These internees did not

require high security hospitalization but were deemed to be too dangerous or unsuitable for admission to a general psychiatric ward or outpatient care (Boers, Vandevelde, Soyez, De Smet, & Ting To, 2011). Conditional release from the internment measure is linked to mandatory treatment under the supervision of a regional court, the 'Commission of the Protection of Society' (CPS). Violent (and other) incidents occurring during the conditional release are reported to the CPS on a regular basis by the MSUs and/or the probation officers.

Participants

The study sample (*n* = 205) was mainly male (91.7%, *n* = 188) and had Belgian nationality (88.7%, *n* = 181)[1]. The mean age at first admission was 36.1 years (*SD* = 11.01, range = 20.8–73.4). The most common diagnoses according to the Diagnostic and Statistical Manual of mental disorders-IV-Text Revision (DSM-IV-TR; American Psychiatric Association, 2000) were personality disorders (76.1%, *n* = 156), substance use disorders (61%, *n* = 125) and psychotic disorders (43.4%, *n* = 89). Levels of comorbidity were high, with 157 internees (76.6%) combining axis I and axis II pathology.

Internment measures were implemented after the following index offenses: violent (75.6%, *n* = 155)[2], property (18%, *n* = 37), drug (2.9%, *n* = 6), non-violent sexual (1.5%, *n* = 3) and other offenses (2%, *n* = 4). Prior contact with the criminal justice services was common, only a small minority (12.2%, *n* = 25) was first offender. The mean number of prior convictions, index offense included, was 6.8 (*SD* = 6.03, range = 1–38).

Sampling procedure

HCR-20 scores coded within one year after MSU admission (*n* = 189) and within one year ahead of discharge (*n* = 132) from the MSU between 2001–2010 were included. Data were gathered by accessing hospital records. Twenty-one HCR-20 assessments at admission and 10 at discharge were excluded because of too many items that were not scored (i.e., four or more missing values in total, two for the historical scale and one for the clinical or risk scale). Further, 17 patients did not spent any time in a less secure setting during the study period, leaving a final sample at discharge of 105 patients, and 168 patients at admission with total HCR-20 scores. Taken together, 38.6% (*n* = 205) of the total population admitted during the study period (*N* = 531) was

assessed either at admission and/or ahead of discharge. Of these 205 patients, a small number (*n* = 24) was assessed on both occasions.

Recidivism data were gathered in 2012-2013 via the Central Criminal Records and the CPS files. Recidivism was evaluated first while still a MSU inpatient (both referring to incidents occurring on the MSU premises and during leave or absconding), and second, after discharge from the MSU. The follow-up period during admission lasted on average almost two years (*M* = 675.4 days, *SD* = 482.27, range = 28–2163). The follow-up period after discharge lasted on average over two years (*M* = 819.7 days, *SD* = 786.99, range = 2–3069), and excluded periods of re-hospitalization in a MSU and detention periods.

Since there were no double clinical scores available to calculate IRR, two researchers with an advanced degree in behavioral sciences scored 26 cases retrospectively, in order to compare the IRR between research and clinical scores. Cases were randomly and proportionally selected according to the number of assessment in each MSU (Bierbeek = 4; Rekem = 9; Zelzate = 13). Assessments were based on file information collected up to the date of original clinical scoring.

Ethical approval was obtained from the Ethics committee of Antwerp University Hospital.

Measurements

HCR-20

The HCR-20 was designed to assess and manage risk of future interpersonal violence and guide risk management plans in settings "in which there is a high proportion of persons with histories of violence, and a strong suggestion of mental illness or personality disorder" (Webster et al., 1997, p. 5). The HCR-20 consists of 20 items (displayed in Table 1) divided across three scales, the Historical (H-), Clinical (C-), and Risk management (R-) scale. Manifestation and relevance of each item is considered, resulting in a summary clinical risk judgment expressed in low, moderate or high risk of violent reoffending. Generally speaking, the more risk factors that are present, the higher the risk of reoffending typically will be. However, the SPJ approach allows for exceptions to this general rule resulting for example in a decision of high risk even when a small number but highly compelling risk factors are present. The final structured professional judgment (referred to in the analyses as 'SPJ') in the present study was sometimes more nuanced, using mixed combinations 'low/moderate' and 'moderate/high'. The category 'low/moderate' was considered

moderate risk and 'moderate/high' was considered high risk. The context of the R-items at discharge ('in' or 'out') was missing in 70.5% of the cases. Furthermore, the authors of the HCR-20 argue that numerical scores can be summed up for research purposes. In the present study, items were sometimes scored using half points. In case items were missing, total and subscale scores (referred to in the analyses as 'numerical scores') were prorated to obtain a score on the same denominator (i.e., 40).

In the present study, the Dutch translation of the HCR-20 was used (Philipse et al., 2000). Research in Dutch clinical samples showed that this version can be reliably scored by trained examiners and scores can meaningfully differentiate between higher and lower risk patients (Hildebrand, Hesper, Spreen, & Nijman, 2005; de Vogel & de Ruiter, 2006). Recently the revised version of the HCR-20, the HCR-20^{V3} (Douglas, Hart, Webster, & Belfrage, 2013) was introduced. Conceptually, this new version is very similar to its predecessor, with the same number of risk factors on each scale, but puts greater emphasis on the usefulness for the individual evaluators. For example, more guidance is provided about individual decision-making, the integration of case material into an explanatory framework and about how summary risk ratings are an important part of constructing useful management plans (Douglas, Hart, et al., 2014). However, version 2 of the HCR-20 was used in the present study since version 3 was not implemented in Flanders at the time of the data collection. The HCR-20 was coded prospectively as part of clinical practice. Scoring was mainly performed by criminologists (professionals with a master degree in Criminology), sometimes in collaboration with a psychologist. Nearly all clinicians were trained in the HCR-20, and all had the necessary professional credentials and were familiar with individual forensic assessments, as the manual prescribes. With respect to the scoring procedure, the consensus method was rarely used, but in most cases, HCR-20 items were completed after team discussions. Additional information on clinical forensic experience in general or familiarity with the HCR-20 in particular, was not available.

Recidivism

Predictive validity was studied using a combined outcome measure of violent recidivism (further referred to as 'recidivism'). First, official sentences for violent offenses were obtained from the Central Criminal Records of the Ministry of Justice. Second, because recidivism data solely based

on new sentences lead to an artificially low base rate in this particular population (e.g., 1.2% and 1.9% for respectively violent recidivism during and after medium security treatment), violent behavior reported to the CPS[3] was added. Violent behavior referred to actual, attempted or threatened interpersonal violence, including sexual violence.

The recidivism rate during MSU treatment was 19.6 % (*n* = 33), and after MSU treatment it was 24.8% (*n* = 26). The recidivism rates of patients with a HCR-20 score did not significantly differ from those of patients without a HCR-20 score both during (19.6% vs. 22.6%, *p* = .44) and after MSU treatment (24.8% vs. 20.4%, *p* = .34).

Analyses

A one-way analysis of variance (ANOVA) or Kruskal Wallis (for non-parametric continuous data) were used to compare mean scores among the three SPJ categories. Chi-square or Fisher exact tests were used to compare the number of recidivists among the three SPJ categories. Paired samples t-tests were used for the changes in normally distributed HCR-20 scores.

IRR was evaluated through a two-way random intraclass correlation coefficient ($ICC_{2,1}$ absolute agreement). Fleiss' (1986) critical values for single measures were used: ICC ≥ .75 = excellent, ICC ≥ .60 = good, ICC ≥ .40 = moderate and ICC < .40 = poor. However, some authors have argued that, in contrast to the standards in basic research, reliability in applied settings ought to be at least .80 if not higher (Heilbrun, 1992; Nunnally, 1978; Rosenthal & Rosnow, 1991).

The correlation between length of stay and the occurrence of recidivism was analyzed using Point-Biserial Correlations in order to determine if this effect needed to be controlled for in subsequent analyses.

Predictive validity was analyzed using seven performance indicators. A global effect size was calculated through the Receiver Operating Characteristic (ROC), producing AUC values. AUC values of .56 or above represent a small effects, .64 or above a moderate effect and .71 or above a large effect (Rice & Harris, 2005), although other authors promote more conservative interpretations with AUC below .60 indicating low accuracy, between .60–.70 marginal accuracy, between .70–.80 modest accuracy, between .80–.90 moderate accuracy, and above .90 high accuracy (Sjöstedt & Grann, 2002). Using the information of a 2 x 2 contingency table, sensitivity (percentage of recidivists who were judged to be at high risk), specificity (percentage of non-

recidivists who were judged to be at low risk), PPV (percentage of participants judged to be at high risk who went on to reoffend), NPV (percentage of low risk individuals who did not go on to reoffend), NND and NSD were calculated to further evaluate predictive accuracy (Singh, 2013). These performance indicators provide information about how accurately the application of a tool identifies high risk ('rule in'; PPV and NND) and low risk ('rule out'; NPV and NSD) individuals. Calculating these measures require a single cut-off threshold. For the numerical scores, this was set at the third quartile. For the SPJ, participants classified as being at high risk were compared with participants classified as low or medium risk. Two-tailed tests were used with a standard significance threshold of $\alpha = .05$.

RESULTS

HCR-20 assessments

Assessment at admission

The mean HCR-20 total score ($n = 168$) was 24.8 ($SD = 5.06$, range = 10.5–36).The mean score on the H-scale was 14.1 ($SD = 3.28$, range = 5–20), on the C-scale 4.8 ($SD = 1.74$, range = 0–8), and on the R-scale 6.0 ($SD = 2.01$, range = 0–10). SPJs were missing for 14 assessments with numerical scores. Nearly one-fifth (18.2%) of the patients ($n = 154$) was classified as low risk, 43.5% as medium and 38.3% as high risk. The mean score for the low risk category was 19.7, for the medium risk category 24.6 and for the high risk category 28.1. This difference of mean scores between the three risk levels was statistically significant ($H(2) = 53.45$, $p = .00$).

In addition, more recidivists were found in the high (27.1%) and medium (16.4%) risk level compared to the low risk level (14.3%) but this difference was not significant ($p = .23$). The relationship between recidivism and length of stay was not significant ($p = .50$).

Assessment at discharge

There was a strong overlap with the first subsample consisting of assessments within one year after admission (43.8%, $n = 46$).

The mean total score was 24.2 ($SD = 5.04$, range = 11.6–35). The mean score on the H-scale was 13.9 ($SD = 3.09$, range = 6.7–20), on the C-scale 4.7 ($SD = 1.68$, range = 0–9) and on the R-scale

5.8 (*SD* = 1.82, range = 2–10). SPJs were missing for four patients. About one-tenth (12.9%) of the patients (*n* = 101) was classified as low risk, 64.4% as medium and 22.8% as high risk. The mean score for the low risk category was 20.0, for the medium risk category 23.4 and for the high risk category 28.5. There was a significant linear trend between mean scores and SPJ categories, indicating that as risk levels increased, mean scores increased proportionately ($F(1, 98) = 31.23$, $p = .00$).

Also, the percentage of recidivists was the highest in the high (39.1%) risk category, followed by the medium (21.5%) and low (7.7%) risk category, but again this difference was not significant (Fisher, $p = .10$). The relationship between recidivism and the length of follow-up after discharge was not significant ($p = .37$). Furthermore, when security level at discharge was taken into account, mean HCR-20 scores differed significantly between high (26.2), medium (24.5) and low (21.5) security levels ($H(2) = 9.51$, $p = .01$). However, this was not the case with the SPJ, where no significant difference was found between risk category and security level at discharge (Fisher, $p = .15$).

Changes of HCR-20 scores

A small subsample (*n* = 24) of the patients with a HCR-20 score were assessed two times – at admission and ahead of discharge. In this subsample, the mean score at admission (*M* = 25.0, *SD* = 5.58, range = 14.0–32.4) did not significantly differ from the mean score at discharge (*M* = 25.4, *SD* = 4.95, range = 15.0–35.0) ($t(23) = -.38$, $p = .71$).

Applicability

Most missing items were found in the H-subscale, but overall, the percentage of missing items was low (Table 1). Item 7, *Psychopathy*, however was omitted in at least half of the assessments. Omitting the item *Psychopathy* has been found to have minimal effect on the HCR-20 and this item has been removed from version 3 (Guy, Douglas, & Hendry, 2010). High item scores were most often found in the historical section. Only for the psychopathy item, a score of 2 was exceptional.

Table 1. Percentage of cases with missing items, distribution of HCR-20 item scores and AUC for the individual items for the two

			Assessment at admission					Assessment at discharge				
			Missing	0	1	2	AUC	Missing	0	1	2	AUC
H	1.	Previous violence	-	4.8	17.3	78.0	.57	-	1.0	20.0	79.0	.51
	2.	Young age at first violent incident	1.8	13.3	43.6	43.0	.53	1.9	12.6	47.6	39.8	.54
	3.	Relationship instability	1.8	9.7	27.3	63.0	.50	-	8.6	35.2	55.2	.54
	4.	Employment problems	4.2	11.8	35.1	51.6	.56	5.7	12.1	44.4	43.4	.46
	5.	Substance use problems	0.6	15.6	13.2	71.3	.54	-	15.2	10.5	74.3	.45
	6.	Major mental illness	0.6	18.6	25.7	55.7	.44	2.9	26.5	22.5	51.0	.48
	7.	Psychopathy	61.3	43.1	49.2	7.7	.59	50.5	48.1	51.9	0	.41
	8.	Early maladjustment	4.2	17.4	19.9	62.7	.59	2.9	18.6	18.6	62.7	.64*
	9.	Personality disorder	1.2	17.5	34.3	60.2	.62*	1.0	21.2	28.8	50.0	.63
	10.	Prior supervision failure	0.6	17.4	22.8	59.9	.49	-	12.4	24.8	62.9	.53
C	1.	Lack of insight	-	1.8	36.9	61.3	.49	-	2.9	37.1	60.0	.38
	2.	Negative attitudes	1.2	57.2	30.7	12.0	.49	-	51.4	40.0	8.6	.58
	3.	Active symptoms of major mental illness	0.6	58.1	25.1	16.8	.44	-	61.9	23.8	14.3	.42
	4.	Impulsivity	-	31.0	41.7	27.4	.64*	-	25.7	50.5	23.8	.71*
	5.	Unresponsive to treatment	-	15.5	51.2	33.3	.52	-	13.3	63.8	22.9	.53
R	1.	Plans lack feasibility	0.6	16.2	56.9	26.9	.51	1.0	12.5	73.1	14.4	.51
	2.	Exposure to destabilizers	-	13.7	42.3	44.0	.46	-	12.4	61.0	26.7	.62
	3.	Lack of personal support	-	19.0	36.3	44.6	.55	-	18.1	37.1	44.8	.51
	4.	Noncompliance with remediation attempts	-	31.0	52.4	16.7	.54	-	25.7	56.2	18.1	.59
	5.	Stress	1.2	7.2	30.7	62.0	.55	-	4.8	43.8	51.4	.62

Note. H = historical items; C = clinical items; R = risk management items; AUC = area under the curve.
* $p < .05$.

Reliability analyses showed mixed rater consistencies. IRR was the highest for the historical scale ($ICC_{2,1}$ = .84). For the total score, IRR was .74. IRR decreased for the C- and R-scale to .64 and .58 respectively. Rater consistency for the SPJ was moderate according to Fleiss' criteria ($ICC_{2,1}$ = .48).[4] According to Heilbrun (1992), only the IRR for the H-scale would be considered acceptable.

Predictive validity

Assessment at admission

Performance indicators for inpatient recidivism are presented in Table 2, while Table 1 provides an overview of AUCs for each item.

Only the individual items *Personality disorder* (H9) and Impulsivity (C4) were able to discriminate recidivists from non-recidivists (significant AUCs with small to moderate effect sizes). AUCs for the total score, subscales and SPJ were non-significant. About two-third of the non-recidivists was classified as low risk, whether the numerical or SPJ approach was used (specificity = 65–71.1%). Half of the recidivists were judged to be at high risk for reoffending when the SPJ was used (sensitivity = 51.6%), while only one-third of the recidivists were judged to be at high risk when actuarial scores were used (sensitivity = 33.3%). The tool was not useful in prospectively predicting who was likely to reoffend (PPV = 22–27.1%), but it did identify low risk patients more accurately (NPV = 81.4–84.2%).

Assessment at discharge

Performance indicators for recidivism after treatment are presented in Table 2, while Table 1 provides an overview of AUCs for each item.

Only the individual items *Early maladjustment* (H8) and *Impulsivity* (C4) were able to discriminate recidivists from non-recidivists (significant AUCs with moderate to large effect sizes). AUCs for the assessments at discharge were non-significant for both the numerical scores and the SPJ. Non-recidivists were correctly classified as being at low risk in more than 70% of the cases (specificity = 76.0–84.8%). One-third of the recidivists were judged to be at high risk when numerical scores as well as SPJ were used (sensitivity = 30.8–37.5%). Low risk individuals were identified with higher accuracy (NPV = 76.9–80.8%) than high risk patients (PPV = 29.6–39.1%). Furthermore, NND was 3 (for both the SPJ and numerical score), meaning that 3 people with a

high risk profile need to be detained in order to prevent one individual from recidivism. Conversely, NSD was 3 and 4 (according to the numerical score and SPJ respectively), meaning that 3 to 4 patients considered to be at low risk (according to the numerical score and SPJ respectively) could be safely discharged prior to an offense.

Table 2. Performance indicators for the numerical scores and structured professional judgment of the HCR-20 in predicting violent recidivism

	At admission (n = 168)	At discharge (n = 105)
Numerical scores		
AUC total (CI)	.60 (.50–.71)	.60 (.48–.72)
AUC H-scale (CI)	.59 (.49–.70)	.57 (.45–.70)
AUC C-scale (CI)	.55 (.43–.66)	.57 (.45–.69)
AUC R-scale (CI)	.54 (.43–.65)	.63 (.52–.74)
Sensitivity	33.3	30.8
Specificity	71.1	76.0
PPV	22	29.6
NPV	81.4	76.9
NND	N/A	3.4
NSD	N/A	3.3
SPJ		
AUC (CI)	.59 (.48–.70)	.63 (.50–.76)
Sensitivity	51.6	37.5
Specificity	65.0	84.8
PPV	27.1	39.1
NPV	84.2	80.8
NND	N/A	2.6
NSD	N/A	4.2

Note. SPJ = structured professional judgment; AUC = area under the curve; PPV = positive predictive value; NPV = negative predictive value; NND = number needed to detain; NSD = number safely discharged; N/A = not applicable.
* $p < .05$.

DISCUSSION

The current study investigated the field validity of the HCR-20, i.e., its relevance to actual clinical risk assessment, by examining IRR between researchers and clinicians and predictive validity of clinical ratings. Predictive validity for violence both during medium security treatment and after treatment was analyzed using seven clinically relevant performance indicators. The main finding of the study was that the HCR-20 was only accurate in identifying low risk individuals.

Mean scores in the current study were somewhat higher than the scores found in medium security samples in the UK, differing about five points on the total scale, three points on the H- and one point on the C- and R-scale (Dolan & Khawaja, 2004; Gray et al., 2004; Gray, Taylor, & Snowden, 2008). Similar scores were found in a medium to high risk sample in the French-speaking part of Belgium (Wallonia; Claix & Pham, 2004) and in forensic psychiatric hospitals in the Netherlands (Mudde, Nijman, van der Hulst, & van den Bout, 2011; de Vogel & de Ruiter, 2006). Although it could be expected that the risk level would decrease as treatment continues and the patient is about to be released, no significant differences could be found in assessments at discharge compared to the ones at admission (n = 22).

IRR for the total score was similar to that of another Belgian study (ICC = .70; Claix & Pham, 2004). Furthermore, compared to the review of Douglas and Reeves (2010), lower IRR was found for the total, C- and R-scale as well as for the SPJ, with only the H-scale having a similar IRR. Likewise, IRR for total scores and SPJs were lower compared to the study by de Vogel and de Ruiter (2004), who also compared clinical with research scores. It should be noted however that in this particular study consistency agreement was analyzed, while the current study used absolute agreement. Also, in the present study, research scores were based on file information only, whereas clinical scores were based on file information including personal contact with the patient. This may have been another explanation of lower IRR for the C-scale and R-scale as compared to the H-scale.

According to Fleiss' (1986) criteria IRR in the current study was considered excellent for the H-scale, good for the total and C- scale and moderate for the R-scale and the SPJ. According to Heilbrun (1992) IRR for all scores would be considered too low for clinical use. The problem of low IRR is particularly relevant in instruments including dynamic risk factors, such as insight in

own problems or stress, and thus for the HCR-20. These items require a great deal of subjective judgment, which may contribute to scoring inconsistencies among raters. The same reasoning could be applied to the SPJ, as clear definitions of low, medium and high risk are lacking.

Based on the (limited) literature examining the predictive validity of the HCR-20 in daily practice, it was hypothesized that the AUC would most likely be lower compared to the performance in carefully planned research designs (Pedersen et al., 2012; Vojt et al., 2013). As field validity studies using performance indicators other than AUC are currently non-existent, no assumptions could be made regarding sensitivity, specificity, PPV or NPV. AUCs were non-significant and lower than those found in research designs in general (Singh et al., 2011) and those found in another Belgian study with a similar population but using research scores (Pham, Ducro, Marghem, & Réveillère, 2005). Comparing to other field validity studies, AUCs were similar to the study of Vojt et al. (2013), but lower than the study of Pedersen et al. (2012).

More specific, only a small number of items showed predictive validity. This is in line with Coid et al. (2011), who demonstrated that the predictive power of the HCR-20 – amongst others – was based on a small number of items. In line with previous research (Coid et al., 2011; Mudde et al., 2011), the item Impulsivity was predictive in the present study, both during and after treatment, and thus seems an interesting dynamic variable to focus treatment on (Coid et al., 2011; Mudde et al., 2011). While impulsivity is broadly defined in the HCR-20 manual, it might be interesting to focus further research on different (behavioral, affective and cognitive) aspects of impulsivity and how these aspects can be reliably assessed. Also, the role of impulsivity is likely to differ depending on the type of violence, i.e., reactive, instrumental and psychotic violence. NPV was relatively higher than PPV indicating that the HCR-20 results are more accurate in making 'rule out' decisions, i.e., identifying individuals at low risk and ready for discharge. This finding is consistent with the suggestion of Fazel et al. (2012) to use risk assessment tools to screen out low risk cases. In contrast, PPV was lower compared to other studies (Singh et al., 2011). Based on a high HCR-20 score unnecessary measures would be taken for three patients in order to prevent one incident in a population with a similar base rate. However, the threshold for predictive accuracy is a legal and human rights issue rather than a clinical question with regards to detainment and deprivation of freedom for patients. It will also be affected by the context

(Buchanan & Leese, 2001). For example, one could argue that false positives are more acceptable in therapeutic conditions compared to imprisonment.

The question remains why accuracy for the prediction of violent recidivism was poor in the current study. One possible explanation for this finding could be that raters were negligent in scoring the HCR-20 or poorly trained. However, in the present study, most clinicians were experienced raters who followed a HCR-20 training course and moderate to high IRR scores indicated that scoring was done as intended.

Another explanation for low predictive accuracy could be that clinicians did not have information on base rates relating to (violent) recidivism since recidivism data were not available at the time of the assessments. Neglecting base rate information is a well-recognized problem in risk assessment, but remains problematic in clinical practice as only one out of five raters makes effective use of this information (Walters et al., 2014).

A final consideration of the poor predictive validity is that the assessment, and thus the knowledge of risk status, resulted in effective risk management. Ultimately, the usefulness of risk assessment can be judged by its ability to contribute to harm reduction. However, if this would be the case, one would expect lower recidivism rates in the group in which the HCR-20 was assessed. This was not the case, since no significant differences were found between patients assessed with the HCR-20 and those without a HCR-20 assessment. Furthermore, although mean HCR-20 scores did significantly correlate with security levels at discharge, SPJs didn't.

While research has focused almost exclusively on predictive validity, it remains unclear whether the use of tools for structured professional judgment actually helps preventing crimes (Wand, 2012). Based on a randomized controlled study investigating this research question in the Netherlands, using another risk assessment tool – the Short Term Assessment of Risk and Treatability (START; Webster, Martin, Brink, Nicholls, & Desmarais, 2009) – it was concluded that the goal of risk prevention was not achieved (Troquete et al., 2013). Equally, in general psychiatry, another violence risk assessment system (the Alert System), although useful at identifying violent patients, did not prevent violent incidents (Kling, Yassi, Smailes, Lovato, & Koehoorn, 2011). Other studies provided more optimistic results on this topic. For example, in two randomized controlled studies performed in general psychiatric units, structured short-term risk assessment using the Brøset Violence Checklist (BVC(-CH); Almvik, Woods, & Rasmussen, 2000) found a reduction in

aggressive incidents and the need for coercive measures (Abderhalden et al., 2008; van de Sande et al., 2011).

Methodological considerations

An important limitation of the current study was the absence of a second HCR-20 assessment by another clinician in order to calculate IRR as scoring consistency improves predictive validity (Hanson & Morton-Bourgon, 2009). Furthermore, consensus ratings are preferred over individual ratings (de Vogel & de Ruiter, 2006). In an attempt to compensate for the lack of a second rater, researchers retrospectively scored 26 files. It should be noted that the research scores were paper-only assessments, whereas clinicians in addition had face-to-face contact with the patients.

Independent of statistical influences, low (or high) reconviction rates, make the outcome difficult to predict, since even highly accurate tests would result in significant error (Szmukler, 2003). Therefore, the current study used a combined outcome measure for recidivism, which yielded a base rate of 19.6% recidivism during treatment and 24.8% recidivism after treatment. It could be argued that including offenses that were not sentenced by a judge are less reliable. However, repeating the analyses with recidivism data based only on the Central Criminal Records (base rates 1.2% and 1.9% respectively) did not change the pattern of the results (analyses available on request).

Furthermore the base rate problem is tempered by the fact that the main goal of SPJ tools is not so much predicting future violence, but reducing this risk. In other words, adequate risk management is associated with low predictive validity. Because the study was based on file information, it was not possible to determine to what extent the ratings actually were taken into account and influenced inpatient risk management and release-decision processes. Likewise, it was not always possible to determine whether assessments at discharge fully took into account the nature of the security level after discharge. The context of R-items at discharge ('in' or 'out') was missing in more than two-thirds of the cases. It is highly likely to influence the accuracy of the R-items.

Finally, data of the three MSUs were not presented separately, so possible differences in the usage of the HCR-20 or the study populations could not be differentiated. However, with the

exception that only one hospital included female patients, the three medium security units used similar inclusion and exclusion criteria.

Clinical implications

As noted by Elbogen (2002), there is a high need of descriptive research "on the implementation and integration of prescriptive risk assessment research" (p. 599), since little is known about when and how prescriptive risk assessment is used in clinical practice. Three identified areas of descriptive research include cue utilization (type of risk factors utilized in practice), clinical reliability and clinical decision making (Elbogen, 2002). The present study confirmed that efforts to increase accuracy of risk assessment are still needed in the field. While the present study used a mean follow-up period of two years, it may be of interest to investigate the predictive validity of the HCR-20 for shorter and longer fixed follow-up periods in future research.

At the clinical level, several recommendations regarding the implementation of risk assessment tools in general and the HCR-20 in particular can be made. First, training is highly important to ensure time effectiveness, reliability and rating consistency (Green et al., 2010). Research investigating the effect of training on risk assessments, found significant improvement in the quality and consistency of the HCR-20 (Reynolds & Miles, 2009). In addition, following initial risk assessment training, continuous peer review processes can minimize drift from item descriptions.

Second, the consensus method of scoring, as proposed by de Vogel and de Ruiter (2006), could further enhance validity. While it might not always be possible to conduct consensus scoring due to time restraints, staff is encouraged to discuss results and preventive measures. Risk assessment can facilitate intra- and inter-professional communication about risk and consistent team responses to danger (Abderhalden et al., 2008). Providing staff with a simple list of clear preventive measures might improve the usage of risk assessment (Abderhalden et al., 2008). Third, as stated in the HCR-20 manual, risk assessments "ought to be made relative to the base rate of violence in a particular population (e.g., low, moderate, or high risk relative to other correctional inmates)" (Webster et al., 1997, p. 7). However, this is easier said than done. First, accurate information on large, non-biased population base rates are often not available to clinicians. Second, even if this information is available, no guidance is provided to clinicians on

how to use this information. Clearly, more needs to be done than simply advising clinicians to use base rate information. In this respect, Walters et al. (2014) suggested to identify cutting scores using information (i.e., means, standard deviations, base rates) derived from a representative sample. They developed an actuarial model as an example of how actuarial prediction can be integrated with base rate information to enhance decision-making accuracy. At least, clinicians should be aware of the fact that in populations with low base rates, neglecting this information is likely to inflate false positive rates.

Surely, a return to unstructured risk assessment and its significant drawbacks is not advised (Roychowdhury & Adshead, 2014). Whether the HCR-20 'works' in a Flemish setting remains to be seen in further studies. In the Netherlands, a specific risk assessment tool for forensic offenders, the HKT-R, was validated nationwide and mandated by the Dutch Ministry of Security and Justice to use in forensic psychiatry regardless of the setting (Spreen, Brand, Ter Horst, & Bogaerts, 2014). While using a Dutch instrument is likely to hamper international comparisons, further research on the predictive validity should examine whether this choice is justified. Recently the revised version of the HCR-20, the HCR-20^{V3}, (Douglas, Hart, Webster, & Belfrage, 2013) was introduced in Flanders. The major aim of this revision was to further improve the applicability in and usefulness for clinical practice. A focus group among clinicians working in medium security units revealed that, in line with Beta-testing by the Dutch translators (de Vogel, van den Broek, & de Vries Robbé, 2014), the first experiences with this new version on the HCR-20 were positive. It was for example argued that the descriptors in the manual were more clear or specific. Although Versions 2 and 3 of the HCR-20 are strongly correlated (.69–.90) (Douglas & Belfrage, 2014), validation studies on this new version in Flanders are also necessary. Furthermore, since evidence of a significant authorship bias was found – especially for peer reviewed articles –, independent research is recommended (Singh, Grann, & Fazel, 2013). Finally, future research should focus on different subgroups of patients and report analyses separately (e.g., women vs. men, personality disordered patients vs. patients with Axis I disorders, intellectually disabled patients vs. patients with normal intellectual abilities), in order to identify potential differences.

CONCLUSION

The current study showed that the overall predictive validity of the HCR-20 in a naturalistic design was low. The tool was mainly effective in identifying low risk individuals. High risk accuracy was poor, meaning that the tool for structured professional judgment failed at accurately predicting who would reoffend. It is interesting to note that there were some items that were predictive, especially impulsivity, which can be examined in future research. The study also revealed that more attention should be paid to implementation issues and continuous peer review.

FOOTNOTES

[1] In one case nationality was missing.

[2] Violent index offenses (*n* = 155) comprised homicide/manslaughter (*n* = 40), sexual assaults (*n* = 12), verbal violence (*n* = 32), and other assaults (*n* = 71). Many of the offenders with non-violent index offenses had a history of convictions for a violent offense. Taken together, 94.6% (*n* = 194) of the population had violently offended either as an index or a prior offense.

[3] Information at the prosecution level was not available in the present study. However, the low sentencing rates in case of reported incidents may be explained due to the mandatory supervision of the CPS, which allows the prosecutor to re-incarcerate the internee within a rather flexible procedure without contradictory debate that is immediately carried out, thereby avoiding the lengthy and lingering procedure of a new trial, which would in most cases result in yet another internment measure.

[4] SPJ was missing for three patients, so ICC was calculated on 23 cases.

REFERENCES

Abderhalden, C., Needham, I., Dassen, T., Halfens, R., Haug, H. J., & Fischer, J. E. (2008). Structured risk assessment and violence in acute psychiatric wards: Randomised controlled trial. *The British Journal of Psychiatry, 193*(1), 44-50. doi: 10.1192/bjp.bp.107.045534

Almvik, R., Woods, P., & Rasmussen, K. (2000). The Brøset Violence Checklist: Sensitivity, specificity, and interrater reliability. *Journal of Interpersonal Violence, 15*(12), 1284-1296

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Boers, A., Vandevelde, S., Soyez, V., Smet, S. De, & To, W.T. (2011). Het zorgaanbod voor geïnterneerden in België. *Panopticon, 2*, 17-38

Buchanan, A., & Leese, M. (2001). Detention of people with dangerous severe personality disorders: A systematic review. *Lancet, 358*(9297), 1955-1959. doi: 10.1016/s0140-6736(01)06962-8

Claix, A., & Pham, T. H. (2004). Evaluation of the HCR-20 Violence Risk Assessment Scheme in a Belgian forensic population. *L'Encéphale, 30*(5), 447-453

Coid, J. W., Yang, M., Ullrich, S., Zhang, T., Sizmur, S., Farrington, D., & Rogers, R. (2011). Most items in structured risk assessment instruments do not predict violence. *The Journal of Forensic Psychiatry & Psychology, 22*(1), 3-21. doi: 10.1080/14789949.2010.495990

DeMatteo, D., Edens, J. F., Galloway, M., Cox, J., Smith, S. T., & Formon, D. (2014). The role and reliability of the Psychopathy Checklist-Revised in U.S. sexually violent predator evaluations: A case law survey. *Law and Human Behavior*, 38(3), 248-255. doi: 10.1037/lhb0000059

Department of Health. (2007). *Best practice in managing risk: Principles and evidence for best practice in the assessment and management of risk to self and others in mental health services*. London: Department of Health.

Dienst Justitiële Inrichtingen. (2014). *Kernset prestatie-indicatoren forensische psychiatrie verslagjaar 2014*. Retrieved from http://www.forensischezorg.nl/files/indicatorengids_verslagjaar_2014_v1.1.pdf

Dolan, M., & Khawaja, A. (2004). The HCR–20 and post-discharge outcome in male patients discharged from medium security in the UK. *Aggressive Behavior, 30*(6), 469-483. doi: 10.1002/ab.20044

Douglas, K. S., & Belfrage, H. (2014). Interrater reliability and concurrent validity of the HCR-20 Version 3. *International Journal of Forensic Mental Health, 13*(2), 130-139. doi: 10.1080/14999013.2014.908429

Douglas, K. S., Hart, S. D., Webster, C., & Belfrage, H. (2013). *HCR-20 (Version 3): Assessing Risk for Violence*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.

Douglas, K. S., Hart, S. D., Webster, C. D., Belfrage, H., Guy, L. S., & Wilson, C. M. (2014). Historical-Clinical-Risk Management-20, Version 3 (HCR-20V3): Development and Overview. *International Journal of Forensic Mental Health*, *13*(2), 93-108. doi: 10.1080/14999013.2014.906519

Douglas, K. S., & Reeves, K. (2010). The HCR-20 violence risk assessment scheme: Overview and re-view of the research. In R. Otto & K. S. Douglas (Eds.), *Handbook of violence risk assessment* (pp. 147-185). Oxford: Routledge/Taylor & Francis.

Douglas, K. S., Shaffer, C., Blanchard, A. J. E., Guy, L. S., Reeves, K. A., & Weir, J. (2014). *HCR-20 violence risk assessment scheme: Overview and annotated bibliography (Current up to January 1, 2014)*. Retrieved from http://kdouglas.wordpress.com/hcr-20/hcr-20-overview-and-annotated-bibliography/

Elbogen, E. (2002). The process of violence risk assessment: A review of descriptive research. *Aggression and Violent Behavior*, *7*, 591–604.

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: Systematic review and meta-analysis. *The British Medical Journal, 345*. doi: 10.1136/bmj.e4692

Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. New York, NY: John Wiley.

Gray, N. S., Snowden, R. J., MacCulloch, S., Phillips, H., Taylor, J., & MacCulloch, M. J. (2004). Relative efficacy of criminological, clinical, and personality measures of future risk of offending in mentally

disordered offenders: A comparative study of HCR-20, PCL:SV, and OGRS. *Journal of Consulting and Clinical Psychology, 72*(3), 523-530. doi: 10.1037/0022-006X.72.3.523

Gray, N. S., Taylor, J., & Snowden, R. J. (2008). Predicting violent reconvictions using the HCR-20. *The British Journal of Psychiatry, 192*(5), 384-387. doi: 10.1192/bjp.bp.107.044065

Green, B., Carroll, A., & Brett, A. (2010). Structured risk assessment in community forensic mental health practice. *Australasian Psychiatry, 18*(6), 538-541. doi: 10.3109/10398562.2010.498513

Guy, L. S., Douglas, K. S., & Hendry, M. C. (2010). The role of psychopathic personality disorder in violence risk assessments using the HCR-20. *Journal of Personality Disorder, 24*(5), 551-580. doi: 10.1521/pedi.2010.24.5.551

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment, 21*(1), 1-21. doi: 10.1037/a0014421

Hare, R. D. (2003). *The Hare Psychopathy Checklist Revised*. Toronto, Ontario, Canada: Multi-Health Systems.

Heilbrun, K. (1992). The role of psychological testing in forensic assessment. *Law and Human Behavior, 16*(3), 257-272. doi: 10.1007/BF01044769

Hildebrand, M., Hesper, B., Spreen, M., & Nijman, H. L. I. (2005). *De waarde van gestructureerde risicotaxatie en van de diagnose psychopathie: Een onderzoek naar de betrouwbaarheid en predictieve validiteit van de HCR-20, HKT-30 en PCL-R*. Utrecht: Expertisecentrum Forensische Psychiatrie.

Hogan, N., & Ennis, L. (2010). Assessing risk for forensic psychiatric inpatient violence: A meta-analysis. *Open Access Journal of Forensic Psychology, 2*, 137-147

Hurducas, C. C., Singh, J. P., Ruiter, C. de, & Petrila, J. (2014). Violence risk assessment tools: A systematic review of surveys. *International Journal of Forensic Mental Health, 13*(3), 181-192. doi: 10.1080/14999013.2014.942923

Kling, R. N., Yassi, A., Smailes, E., Lovato, C. Y., & Koehoorn, M. (2011). Evaluation of a violence risk assessment system (the Alert System) for reducing violence in an acute hospital: A before and after study. *International Journal of Nursing Studies, 48*(5), 534-539. doi: 10.1016/j.ijnurstu.2010.10.006

Ministry of Health. (1998). *Guidelines for clinical risk assessment and management in mental health services*. New Zealand: Ministry of Health.

Mudde, N., Nijman, H., Hulst, W. van der, & Bout, J. van den (2011). Het voorspellen van agressie tijdens de behandeling van forensisch psychiatrische patiënten aan de hand van de HCR-20 [Predicting aggression during the treatment of forensic psychiatric patients by means of the HCR-20]. *Tijdschrift voor Psychiatrie, 53*(10), 705-713

Neal, T. M. S., Miller, S. L., & Shealy, R. C. (2015). A field study of a comprehensive violence risk assessment battery. *Criminal Justice and Behavior, 42*(9), 952-968. doi: 10.1177/0093854815572252

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

O'Shea, L. E., Mitchell, A. E., Picchioni, M. M., & Dickens, G. L. (2013). Moderators of the predictive efficacy of the Historical, Clinical and Risk Management-20 for aggression in psychiatric facilities: Systematic review and meta-analysis. *Aggression and Violent Behavior, 18*(2), 255-270. doi: 10.1016/j.avb.2012.11.016

Pedersen, L., Ramussen, K., & Elsass, P. (2012). HCR-20 violence risk assessments as a guide for treating and managing violence risk in a forensic psychiatric setting. *Psychology, Crime & Law, 18*(8), 733-743. doi: 10.1080/1068316X.2010.548814

Penney, S. R., McMaster, R., & Wilkie, T. (2014). Multirater reliability of the historical, clinical, and risk management-20. *Assessment, 21*(1), 15-27. doi: 10.1177/1073191113514107

Pham, T. H., Ducro, C., Marghem, B., & Réveillère, C. (2005). Évaluation du risque de récidive au sein d'une population de délinquants incarcérés ou internés en Belgique francophone [Prediction of recidivism among prison inmates and forensic patients in Belgium]. *Annales Médico-psychologiques, revue psychiatrique, 163*(10), 842-845. doi: 10.1016/j.amp.2005.09.013

Philipse, M., Ruiter, C. de, Hildebrand, M., & Bouman, Y. (2000). *HCR-20. Beoordelen van het risico van gewelddadig gedrag* (Versie 2). Nijmegen/Utrecht: Pompestichting/Dr. Henri van der Hoeven Stichting.

Reynolds, K., & Miles, H. L. (2009). The effect of training on the quality of HCR-20 violence risk assessments in forensic secure services. *The Journal of Forensic Psychiatry & Psychology, 20*(3), 473-480. doi: 10.1080/14789940802638366

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's *d*, and *r*. *Law and Human Behavior, 29*(5), 615–620. doi: 10.1007/s10979-005-6832-7

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York, NY: McGraw-Hill.

Roychowdhury, A., & Adshead, G. (2014). Violence risk assessment as a medical intervention: Ethical tensions. *The Psychiatric Bulletin, 38*(2), 75-82. doi: 10.1192/pb.bp.113.043315

Rufino, K. A., Boccaccini, M. T., & Guy, L. S. (2011). Scoring subjectivity and item performance on measures used to assess violence risk: The PCL-R and HCR-20 as exemplars. *Assessment, 18*(4), 453-463. doi: 10.1177/1073191110378482

Sande, R. van de, Nijman, H. L., Noorthoorn, E. O., Wierdsma, A. I., Hellendoorn, E., Staak, C. van der, & Mulder, C. L. (2011). Aggression and seclusion on acute psychiatric wards: Effect of short-term risk assessment. *British Journal of Psychiatry, 199*(6), 473-478. doi: 10.1192/bjp.bp.111.095141

Singh, J. P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences & the Law, 31*(1), 8-22. doi: 10.1002/bsl.2052

Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., . . . Otto, R. K. (2014). International perspectives on the practical application of violence risk assessment: A global survey of 44 countries. *International Journal of Forensic Mental Health, 13*(3), 193-206. doi: 10.1080/14999013.2014.922141

Singh, J. P., & Fazel, S. (2010). Forensic risk assessment: A metareview. *Criminal Justice and Behavior, 37*(9), 965-988. doi: 10.1177/0093854810374274

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499-513. doi: 10.1016/j.cpr.2010.11.009

Singh, J. P., Grann, M., & Fazel, S. (2013). Authorship bias in violence risk assessment? A systematic review and meta-analysis. *PLoS One, 8*(9), e72484. doi: 10.1371/journal.pone.0072484

Sjöstedt, G., & Grann, M. (2002). Risk assessment: What is being predicted by actuarial prediction instruments? *International Journal of Forensic Mental Health, 1*(2), 179-183. doi: 10.1080/14999013.2002.10471172

Spreen, M., Brand, E., Horst, P. Ter, & Bogaerts, S. (2014). *Handleiding en methodologische verantwoording HKT-R, Historisch, Klinische en Toekomstige - Revisie*. Groningen: Dr. S. van Mesdag.

Szmukler, G. (2003). Risk assessment 'numbers' and 'values'. *Psychiatric Bulletin, 27*(6), 205-207. doi: 10.1192/pb.27.6.205

Troquete, N. A., Brink, R. H. van den, Beintema, H., Mulder, T., Os, T. W. van, Schoevers, R. A., & Wiersma, D. (2013). Risk assessment and shared care planning in out-patient forensic psychiatry: Cluster randomised controlled trial. *The British Journal of Psychiatry, 202*(5), 365-371. doi: 10.1192/bjp.bp.112.113043

Vogel, V. de, Broek, E. van den, & Vries Robbé, M. de (2014). The use of the HCR-20V3 in Dutch forensic psychiatric practice. *International Journal of Forensic Mental Health*, *13*(2), 109-121. doi: 10.1080/14999013.2014.906518

Vogel, V. de, & Ruiter, C. de (2004). Differences between clinicians and researchers in assessing risk of violence in forensic psychiatric patients. *The Journal of Forensic Psychiatry & Psychology, 15*(1), 145-164. doi: 10.1080/14789940410001655916

Vogel, V. de, & Ruiter, C. de (2006). Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime & Law, 12*(3), 321-336. doi: 10.1080/10683160600569029

Vogel, V. de, Ruiter, C. de, Hildebrand, M., Bos, B., & Ven, P van de. (2004). Type of discharge and risk of recidivism measured by the HCR-20: A retrospective study in a Dutch sample of treated forensic psychiatric patients. *International Journal of Forensic Mental Health*, *3*(2), 149-165. doi: 10.1080/14999013.2004.10471204

Vojt, G., Thomson, L. D. G., & Marshall, L. A. (2013). The predictive validity of the HCR-20 following clinical implementation: Does it work in practice? *The Journal of Forensic Psychiatry & Psychology, 24*(3), 371-385. doi: 10.1080/14789949.2013.800894

Walters, G. D., Kroner, D. G., & DeMatteo, D. (2014). The impact of base rate utilization and clinical experience on the accuracy of judgments made with the HCR-20. *Journal of Forensic Psychology Practice, 14*(4), 288-301. doi: 10.1080/15228932.2014.941726

Wand, T. (2012). Investigating the evidence for the effectiveness of risk assessment in mental health care. *Issues in Mental Health Nursing, 33*(1), 2-7. doi: 10.3109/01612840.2011.616984

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence* (Version 2). Burnaby, BC: Simon Fraser University.

Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Desmarais, S. (2009). *Manual for the Short-Term Assessment of Risk and Treatability (START)*. Port Coquitlam, BC: Forensic Psychiatric Services Commission and St. Joseph's Healthcare.

Wilson, C. M., Desmarais, S. L., Nicholls, T. L., Hart, S. D., & Brink, J. (2013). Predictive validity of dynamic factors: Assessing violence risk in forensic psychiatric inpatients. *Law and Human Behavior, 37*(6), 377-388. doi: 10.1037/lhb0000025

PCL-R field validity in prison and hospital settings

Law and Human Behavior, in press

9

Jeandarme, I., Edens, J., Habets, P., Bruckers, L., Oei, T. I., & Bogaerts, S.

ABSTRACT

Recent field studies have questioned the interrater reliability (IRR) and predictive validity regarding (violent) recidivism of the Psychopathy Checklist-Revised (PCL-R). Using a forensic psychiatric sample, the current study investigated discrepancies in scoring between hospital and prison settings, as well as differences in predictive validity across these 2 settings. PCL-R information was collected from prison and hospital files, resulting in 224 PCL-R total scores and 74 double scores. When examining repeated measurements, large individual differences were found together with an intraclass correlation coefficient ($ICC_{A,1}$) of .42 for the total score. Discrepant results were found for Factor 2, with repeated scores within the same setting having an $ICC_{A,1}$ of .28 versus an $ICC_{A,1}$ of .57 for repeated scores between settings. However, areas under the curve (AUCs) from receiver operating characteristic (ROC) analyses for total, factor and facet scores did not differ between settings. For the whole sample, Factor 2 scores marginally predicted violent and general recidivism after 2 years (AUC = .62 and .63), whereas Factor 1 did not predict (violent) recidivism. Consistent with recent studies from other countries, these results suggest inadequate field reliability and validity in prison and hospital settings in Flanders (Belgium).

INTRODUCTION

The Psychopathy Checklist Revised (PCL-R; Hare, 2003) is an extensively used and researched instrument for diagnosing psychopathy. Early factor analyses suggested that the PCL-R consisted of two factors: Factor 1 representing the interpersonal and affective component and Factor 2 capturing the socially deviant and behavioral aspects (Hare, Clark, Grann, & Thornton, 2000). Later, Hare (2003) argued for the existence of a superordinate factor of psychopathy, underpinned by two factors (interpersonal/ affective and social deviance) and four facets (interpersonal, affective, lifestyle, and antisocial; cf. Cooke, Michie, & Hart, 2006). The PCL-R is also frequently introduced in the legal arena to inform violence risk assessment (DeMatteo, Edens, Galloway, Cox, Smith, Koller, et al., 2014), either in isolation or included as an important component within risk assessment instruments, such as the Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice, & Cormier, 2006), and the Historical Clinical Risk Management–20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997). It should be noted that the revised version of the HCR-20 (HCRV3; Douglas, Hart, Webster, & Belfrage, 2013) no longer requires a PCL-R and that in the revised version of the VRAG, the PCL-R was replaced by Facet 4 of the PCL-R (VRAG-R; Rice, Harris, & Lang, 2013). Over the years, research from controlled research settings has shown good psychometric properties for the PCL-R and moderate association with criminal recidivism in a variety of cohorts, including forensic psychiatric patients (Singh, Grann, & Fazel, 2011; Yang, Wong, & Coid, 2010).

Although scoring the PCL-R requires at least some subjective judgment, strong interrater reliability (IRR) with good to excellent intraclass correlation coefficients (ICCs) for the total score (.86 to .94), Factor 1 (.69 to .95) and Factor 2 (.74 to .94) have been reported in early validation studies as well as in independent controlled research (e.g., Cooke, Hart, & Michie, 2004; Gacono & Hutton, 1994; Hare, 2003; Ismail & Looman, 2016; Kroner & Mills, 2001; Laurell & Daderman, 2007; Porter, Woodworth, Earle, Drugge, & Boer, 2003). The ICCs for facet scores have ranged from .67 for Facet 2 to .93 for Facet 4 (Hare, 2003; Ismail & Looman, 2016). In addition, when adequate archival material is available to score the PCL-R items, good to excellent agreement between research and clinical ratings has been reported for the total score in some studies

(Declercq, Willemsen, Audenaert, & Verhaeghe, 2012; Grann, Langstrom, Tengstrom, & Stalenheim, 1998; Harris, Rice, & Cormier, 2013).

What exactly constitutes an acceptable level of interrater reliability for a rating scale such as the PCL-R is open to debate. Noted authorities have argued that, in contrast to suggested standards in basic research (i.e., ICC ≥ .60; Fleiss, 1986), in applied clinical and forensic settings reliability ought to be at least .80 if not higher (Heilbrun, 1992; Nunnally, 1978; Rosenthal & Rosnow, 1991) because of the real world consequences of rater error. This would seem especially true for an instrument such as the PCL-R, where scores are used as evidence in many high stakes contexts (e.g., capital punishment, indeterminate civil commitment) to influence legal decision-making (DeMatteo, Edens, Galloway, Cox, Smith, & Formon, 2014). However, values for the PCL-R total score reported in the manual (Hare, 2003) and in other controlled research studies typically have approximated these recommended levels of IRR for applied settings (i.e., ICC ≥ .80).

Field reliability studies

In contrast to reliability statistics described in the PCL-R manual and in most controlled research studies, PCL-R scores that raters assign as part of routine clinical or forensic practice, appear to be considerably less reliable, based on recent field studies conducted mostly in North America. Research focusing on rater differences in applied settings has focused primarily on two non-mutually exclusive explanations for these differences: adversarial allegiance and individual differences across raters.

Adversarial allegiance

There is growing evidence (e.g., DeMatteo, Edens, Galloway, Cox, Smith & Formon, 2014, Edens, Cox, Smith, DeMatteo, & Sorman, 2015; Lloyd, Clark, & Forth, 2010; Murrie, Boccaccini, Johnson, & Janke, 2008; Murrie et al., 2009) that PCL-R scoring is affected by the evaluation context, with adversarial settings such as contested criminal or civil commitment cases producing scores that diverge much more so than would be expected based on the ICC statistics reported in the professional manual (Hare, 2003). It has been argued that this is at least in part due to adversarial allegiance in legal proceedings, where opposing experts are retained by different sides of a case and the subsequent scores they produce are biased by a pull to affiliate with the

legal party who retained the examiner. Much of this research has been conducted in sexually violent predator (SVP) trials in the U.S., where the first studies (Murrie et al., 2008; Murrie et al., 2009) reported IRRs for PCL-R total scores ranging from .39 to .42 ($ICC_{A,1}$) between opposing evaluators (prosecution vs. defense). A more recent case law review (DeMatteo, Edens, Galloway, Cox, Smith & Formon, 2014) reported a somewhat higher value ($ICC_{A,1}$ = .68 and kappa = .46) for U.S. sex offenders (*n* = 14), although still below acceptable standards for forensic practice.

The above-noted studies were based on relatively small U.S. SVP samples, but Edens, Cox, Smith, DeMatteo, and Sörman (2015) came to similar conclusions when comparing Crown versus defense experts in 55 Canadian Dangerous and Long Term Offender (DLTO) evaluations ($ICC_{A,1}$ = .54; see also Lloyd et al., 2010). In all of the above-noted studies, average scores provided by prosecution-retained experts have been significantly higher compared to experts retained by the defense, with mean differences generally ranging between approximately 5 and 8 points (DeMatteo, Edens, Galloway, Cox, Smith & Formon, 2014).

Due to the uncontrolled nature of field studies (i.e., lack of randomization), these results in principle could be explained by something other than adversarial allegiance, but the same general pattern of results recently was demonstrated in a study using a true experimental manipulation. Murrie, Boccaccini, Guarnera, and Rufino (2013) randomly assigned mental health professionals who had been trained on the PCL-R to conditions in which they were instructed that the attorney with whom they were consulting on SVP cases was employed by either the prosecution or the defense. Those randomly assigned to the prosecution condition subsequently provided significantly higher PCL-R scores compared with those who believed they had been consulting with the defense, with Cohen's *d* values ranging between .55 and .85 for three of the four cases2 being evaluated. Murrie et al. (2013) also examined the reliability of the Static-99R, a risk assessment measure that includes 10 items that mainly refer to factual information from the past and requires less clinical judgment than the PCL-R. Notably, although there were some trends toward allegiance effects on this instrument as well, the differences were much smaller in magnitude across the cases (i.e., *d* values ranged from .14 to .44).

Evaluator scoring differences

Although adversarial allegiance appears to be a significant source of concern in contested legal proceedings, the relatively poor reliability of PCL-R scores in applied settings cannot be explained solely as a function of mental health examiners being retained by competing sides of the same case. Research into this topic suggests that individual examiners who conduct numerous PCL-R evaluations in the field may show large differences in the average scores that they tend to provide for the individuals they are assessing. For example, Boccaccini, Turner, and Murrie (2008) identified two examiners who between them had conducted 100 separate SVP evaluations for the state of Texas who differed in their average reported PCL-R scores by almost 10 points (Cohen's *d* = 1.27). Although these two raters perhaps were somewhat atypical, more systematic reliability analyses examining the influence of the individual rater on PCL-R scores in these SVP cases indicated that over 30% of the variance in total scores was attributable to which particular examiner had been assigned to the case. In subsequent analyses of an overlapping but somewhat larger sample (*n* > 550), Boccaccini, Murrie, Rufino, and Gardner (2014) reported a similar result for total scores (32%) and also estimated that approximately one quarter of the variance in Factor 1 and 2 scores was attributable to examiner effects.

Given the significant role that individual raters seem to play in the PCL-R scores they provide, it is not surprising that reliability across examiners is not particularly high even if those examiners are retained by the same side in a given case. In comparing evaluators who were retained by the same side in SVP cases, Boccaccini et al. (2008) found an $ICC_{A,1}$ of only .47 for total PCL-R scores (*n* = 22), whereas Miller, Kimonis, Otto, Kline, and Wasserman (2012) reported somewhat higher ICCs (ICC1 for total PCL-R score = .60) in a much larger SVP sample (*n* = 313) from Florida. Similarly, Edens et al. (2015) reported that ICCs tended to be relatively low in their Canadian DLTO sample (most in the .5 to .6 range) regardless of whether the examiners had been retained by the same side of a case, were court-appointed, or were retained by opposing parties. Among published field studies to date, Sturup et al. (2014) obtained the highest level of agreement for the total score ($ICC_{A,1}$ = .70) among 27 life sentenced offenders in Sweden who had been evaluated by a nonpartisan government agency. Unlike other field studies, however, both of the PCL-R scores examined in their analyses were derived from the consensus ratings of separate

multidisciplinary teams (with each team typically consisting of a forensic social investigator, psychologist, and forensic psychiatrist).

Although field reliability studies comparing examiners retained by the same side of a case have focused primarily on the PCL-R total score, some of this research has been able to investigate this topic at the factor, facet, and even item level. In the largest study to date (with *n*s ranging from 147 to 154), Miller et al. (2012) showed in their Florida SVP sample that Factor 1 scores were considerably less reliable (ICC_1 = .48) than Factor 2 scores (ICC_1 = .72) and Facet 2 (ICC_1 = .39) was scored much less reliable than Facet 4 (ICC_1 = .75). More recently, Sturup et al. (2014) also reported lower reliability on the first three facets ($ICC_{A,1}$ = .54 – .60), compared with Facet 4 ($ICC_{A,1}$ = .90) in their life-sentenced Swedish sample. This general pattern of results is consistent with both other field research (Edens, Boccaccini, & Johnson, 2010) and nonfield-based research investigations (Hare, 2003) suggesting that Factor 1 and its two facets are typically significantly less reliable than Factor 2 and its two facets.

In terms of item-level reliability, individual ICCs have ranged widely both within and across studies (e.g., from .09 to .73 in the Miller et al., 2012 Florida SVP sample and from .23 to .80 in the Sturup et al., 2014 Swedish sample), with Facet 4 items (e.g., juvenile delinquency, criminal versatility) typically being the most stable across raters. In the Miller et al. (2012) sample, individual items from Factor 1 seemed to be especially problematic, such as Item 7 (*shallow affect*, ICC_1 = .09), Item 5 (*conning/manipulative*, ICC1 = .26), and Item 4 (*pathological lying*, ICC1 = .29). Similarly, in the Sturup et al. (2014) sample, rater agreement was particularly problematic for Item 1 (*glibness*, $ICC_{A,1}$ = .31), Item 8 (*callous*, $ICC_{A,1}$ = .36), and Item 5 (*conning/manipulative*, $ICC_{A,1}$ = .37), although the worst-performing item overall was from the lifestyle facet (*impulsivity*, $ICC_{A,1}$ = .23). Although the abovementioned studies indicate that individual examiners account for some of the difference in scoring, few studies have investigated exactly *why* such differences occur. One study by Miller, Rufino, Boccaccini, Jackson, and Murrie (2011) identified personality traits of the evaluator as one potential explanation for rater disagreement. Evaluators higher in *agreeableness* (assessed with the NEO Personality Inventory—Revised; Costa & McCrae, 1992) assigned lower PCL-R scores on the interpersonal facet (*compassion hypothesis*) and more extraverted people with higher excitement seeking tended to provide lower scores on the impulsivity facet (*normalization hypothesis*). Other potential explanations relate to individual

differences in the level of experience raters have in scoring the PCL-R. Research has suggested that more experienced evaluators tend to assign lower scores than less experienced evaluators (Rufino, Boccaccini, Hawes, & Murrie, 2012). Furthermore, level of background training and education could influence scoring. In some studies, scoring was mostly performed by psychologists with a master's degree and to a lesser extent psychiatrists (e.g., Miller et al., 2012), whereas some studies included mostly psychologists with a doctoral degree (e.g., Boccaccini, Turner, & Murrie, 2008; Murrie et al., 2013) or combinations of doctoral level psychologists and psychiatrists (DeMatteo, Edens, Galloway, Cox, Smith, & Formon, 2014; Edens et al., 2015). It could be argued that highly trained researchers and clinicians with a doctoral degree might exhibit greater skill and objectivity than is typically the case among clinicians in general. There is some evidence that doctoral-level examiners who have completed some formalized PCL-R training may produce more reliable scores (Boccaccini et al., 2014) and more prolific evaluators may produce scores that demonstrate somewhat better predictive validity as well (Murrie, Boccaccini, Caperton, & Rufino, 2012).

Predictive validity

The association between the PCL-R and future inpatient and outpatient offending has received considerable attention in the literature. In meta-analyses of the predictive efficacy of risk assessment schemes, mostly moderate effect sizes have been reported for the PCL-R total score (*d* = .55, Leistico, Salekin, DeCoster, & Rogers, 2008; AUC = .66, Singh et al., 2011; AUC = .65, Yang et al., 2010). Also, research has shown differences in predictive validity between factors, with Factor 2 scores demonstrating stronger effect sizes for violent recidivism (AUC = .67, Yang et al., 2010), and general recidivism (Walters, 2003). Yang et al., (2010) failed to identify a significant effect for Factor 1 for men, although a small effect emerged for women. Although meta-analyses have strongly questioned the relevance of Factor 1 scores to risk assessment, we should note that a few individual studies have highlighted the potential importance of Factor 1. For example, Laurell, Belfrage, and Hellstrom (2010) reported that Factor 1 scores were related to instrumentality in the index crime and to the degree of victim injury. At the facet level, most research (e.g., Walters, Knight, Grann, & Dahle, 2008) has indicated that Facet 4 is a stronger predictor of general and violent recidivism than the remaining facets, which is consistent with

research suggesting that Facet 4 is also the subscale that is most reliably scored by raters in both basic and field research.

Leistico et., (2008) examined several potential moderator variables that might impact the magnitude of the effect sizes (e.g., types of settings in which predictive studies were performed, racial and gender composition of samples). In their meta-analysis, Factor 2 scores from patients in forensic or civil hospitals had better predictive validity than scores from participants in prison settings. Also, larger effect sizes for total and Factor 2 PCL-R scores were reported when the study sample contained larger percentages of Caucasian participants. Total and Factor 1 scores also explained future antisocial behavior better in samples that included more female participants, although such a result could simply be an artifact of including mixed gender samples (i.e., women being typically being a lower risk for recidivism and also generally receiving lower PCL-R scores than men, resulting in inflated effect sizes).

In terms of predictive validity across basic and applied research settings, one would expect that lower rater agreement would imply lower predictive validity in applied settings. Supporting this general position, in a meta-analysis of sexual reoffending by Hawes, Boccaccini, and Murrie (2013), the predictive validity for sexual recidivism differed between researchers and clinicians, with stronger effect sizes reported for research purpose (d = 0.44) compared with those calculated for clinical use (d = 0.28). In a similar vein, a very recent field study in a forensic hospital using archival data from 230 case files reported broad null predictive validity findings for various outcome measures, with none of the subscales predicting better than chance (Neal, Miller, & Shealy, 2015). These findings notwithstanding, lower ICC values across raters do not mandate that predictive validity in the same setting will be uniformly lower. For example, Boccaccini, Turner, Murrie, and Rufino (2012) found in SVP trials that total scores from opposing evaluators were moderate to strong predictors of misconduct among 57 sex offenders after their civil commitment (AUC = .71–.77) despite relatively low IRR across examiners. It seems there would be at least two potential reasons for such findings. First, it is possible that individual evaluators tend to rank-order offenders in the same general manner, even though they might demonstrate large mean differences in scores across these cases (Harris, Boccaccini, & Murrie, 2015). Second, mixing essentially *more accurate* and *less accurate* raters to assess the same cases would likely result in lower IRR even though the "good" raters were providing PCL-R scores that

would be more valid predictors of risk. As a potential example of such an effect, Murrie et al. (2012) investigated postrelease arrest rates in a large sample of sex offenders (*n* = 333) screened for SVP (but subsequently not committed). ICCs have been poor in the various field studies previously reported by Murrie et al. (2012) and, not surprisingly, none of the AUC values for recidivism were significant for the total sample in this particular study. PCL-R scores from a small subset of forensic evaluators (*n* = 3), however, did predict violent recidivism better than scores from the other 15 examiners. Notably, these examiners were three of the four most prolific in the sample, having conducted a relatively large minority of the total number of SVP assessments.

In summary, the small but accumulating body of literature suggests considerably attenuated reliability and predictive validity when the PCL-R is used in applied forensic settings. Although the first field studies were based on relatively small samples from one jurisdiction (Texas) involving a specific population (sex offenders), other studies have provided further evidence of lower reliability based on larger samples in other U.S. jurisdictions (DeMatteo, Edens, Galloway, Cox, Smith, Koller, et al., 2014; Levenson, 2004; Miller et al., 2012; Neal et al., 2015) as well as in Canadian and European samples (Edens et al., 2015; Sturup et al., 2014). The potential adverse implications of these field validity findings should not be taken lightly, given concerns about the stigmatizing effects of the *psychopath* label (e.g., Bersoff, 2002; Edens, Davis, Fernandez Smith, & Guy, 2013; Lloyd et al., 2010; Wayland & O'Brien, 2014) and particularly Factor 1-type characteristics (Edens, Colwell, Desforges, & Fernandez, 2005; Edens et al., 2013; Sundby, 1997) on sentencing decisions trials and legal decision-making. Based on the field reliability/validity literature and the stigmatization research, some commentators (Edens, Petrila, & Kelley, in press) have gone so far as to suggest that examiners should not introduce Factor 1 scores from the PCL-R into legal proceedings because of their limited probative value and their potential to introduce undue prejudice toward defendants. However, the total number of field validity studies is relatively limited (particularly in relation to examinations of factor scores), and sometimes based on small samples (e.g., *n* = 14 to 55; DeMatteo, Edens, Galloway, Cox, Smith, Koller, et al., 2014; Edens et al., 2015; Lloyd et al., 2010; Murrie et al., 2008; Murrie et al., 2009). As such, examining the reproducibility of these results, particularly in jurisdictions outside of North American and with samples other than sex offenders, is particularly important.

The current study

The current study extends to the existing research concerning the field reliability and validity of the PCL-R by examining this topic in a relatively large sample of Belgian offenders found not guilty by reason of insanity (NGRI) (in Belgium referred to as internees) who were classified within a medium security risk level. PCL-R assessments were conducted while the patients resided in prison and/or in hospital. For a large minority of the sample, multiple scores were available for analysis. Although this is a largely descriptive study, based on the literature reviewed earlier, we hypothesized that reliability statistics would be significantly poorer than those reported in the PCL-R manual (Hare, 2003) and that predictive validity would be substantially weaker than what has been reported in meta-analyses (e.g., Leistico et al., 2008; Yang et al., 2010) and more consistent with other recent field studies (e.g., Murrie et al., 2012; Neal et al., 2015) demonstrating very modest effect sizes.

Also, given differences in the staff conducting these assessments across settings (e.g., education), we anticipated that rater agreement would be lower across settings (hospital vs. prison), as compared with within settings (hospital vs. hospital and prison vs. prison).

Although not adversarial allegiance per se, we expected that the study might show evidence of some kind of partisanship or contextual pressures that might impact scores across these settings. For example, given systemic pressures within the prison system, it could be argued that prison evaluators might feel compelled to score an examinee lower on the PCL-R than in forensic institutions in order to facilitate transfer of forensic patients from prison to hospital. Furthermore, prison evaluators might have a higher threshold for identifying traits as psychopathic, due to different prototypical views of psychopathy compared with hospital evaluators who work primarily with patients who tend to be relatively less psychopathic than the average prisoner. However, there are at least some contextual factors such as therapeutic relationships that might result in hospital staff providing somewhat lower psychopathy ratings. For example, researchers investigating HCR-20 assessments have found that group leaders conducting the daily supervision and spending most of their time with the patients tended to give lower HCR-20 scores (de Vogel, de Ruiter, Hildebrand, Bos, & van de Ven, 2004).

METHOD

Setting

This study is part of a larger study investigating recidivism in 531 offenders found NGRI and referred to medium security treatment in one of the three Flemish medium security forensic units (MSU; Jeandarme, Habets, Oei, & Bogaerts, 2016). On the basis of their mental state (determined to be either completely irresponsible or severely diminished in responsibility for the committed crime) and dangerousness at the time of the trial, these offenders were initially subjected to a preventive *internment* measure rather than being convicted. However, due to a shortage of forensic psychiatric units, NGRI offenders often reside in prison until a treatment setting is found. This was the case for the majority of the sample (96.9%, *n* = 217). The decision to transfer the NGRI offenders to a MSU for mandatory treatment was determined by a dedicated legal authority, the Commission of the Protection of Society (CPS), which is chaired by a judge. Because the initial treatment in MSU and subsequent treatment in less secure units or the community was delivered under the judicial authority of the CPS, incidents that occurred during treatment, were reported on a regular basis to the CPS by the judicial assistants. The CPS decided on the duration of the internment as well as any revocation of conditional release, (i.e., return to prison) if required. Reported incidents typically comprised general misconduct such as noncompliance to treatment or absconding from treatment but also included new crime-related incidents (e.g., new assaults). In a large minority (42.2%, *n* = 224) of MSU patient records reviewed for this archival study, a PCL-R score was reported. In Belgium, the use of the PCL-R was (and still is) not mandatory in the routine assessment of forensic psychiatric patients, but based on an informal survey of MSU staff conducted by the authors, clinicians who work on these units reported finding the instrument useful in certain cases, mainly for diagnostic purposes or risk assessment and management issues. Referral indicators for performing a PCL-R evaluation included mainly treatment amenability, lengthy criminal history, antisocial personality disorder, history of predatory violence, and violence potential. In prison settings, a PCL-R assessment was conducted as part of the conditional release procedure, but only if sufficient collateral information was available to appropriately rate the items, as recommended by Hare (2003).

Multiple PCL-R scores were located in 21.4% (*n* = 48) of the files. The mean time difference between measurements was 3.2 years (*M* = 1166.5 days, *SD* = 994.5 days, range = 6–4,741). Typically, a second assessment was performed when a substantial period of (conditional) release occurred after the first scoring. Sometimes, staff were simply not aware of the first score. Being a field study, not all evaluators were trained in PCL-R scoring by the same supervisors (e.g., some were trained in Belgium, some in the Netherlands), but all evaluators who scored the PCL-R had a master's or MD degree and were trained in a manner consistent with the procedures outlined in the test manual (Hare, 2003).

Participants

Most participants (*N* = 224) were male (98.7%). Mean age at the time of the PCL-R assessment was 35.3 years (*SD* = 8.9; range = 18.6 –59.8) (31 missing cases). Participants had on average 6.9 prior convictions for a general offense such as theft or drug offenses (*Mdn* = 6.0, *SD* = 6.0, range = 1–38) and 2.5 prior convictions for a violent crime (*SD* = 1.9, range = 0–8). Index offenses were primarily (79.5%) violent in nature (including sex offenses), whereas 16.5% were property crimes, 1.3% were hands-off sexual crimes, 1.3% were drug crimes, and 1.3% were categorized as other crimes (mainly traffic-related, such as driving without insurance/license). Average age at first offense was 23.3 years (*SD* = 7.0, range = 10.2–53.6). The majority of the participants had Belgian nationality (91.1%) and had been previously admitted to a psychiatric hospital (80.4%) (five missing cases). Two participants died during the study (one due to illness and one due to suicide). Participants exhibited the following Axis I diagnoses: substance-related disorders (58.9%), psychotic disorders (37.5%), other disorders (25.4%), no diagnosis (7.6%), developmental disorders (6.3%), mood disorders (5.8%), paraphilia (4.0%), panic disorders (1.8%), and cognitive disorders (0.4%). Axis II diagnoses established in the participants were: Cluster B personality disorders (50.9%), no or deferred diagnoses (17.0%), personality disorders NOS (16.1%), Cluster A personality disorders (7.6%), and Cluster C personality disorders (3.6%) and intellectual disability (12.5%). Average IQ score was 81.8 (*SD* = 15.8, range = 48–138). The mean HCR-20 score reported in archival records was 25.3 (*SD* = 5.2, range = 11–36). Comparing the group with and without a PCL-R score in the current study, the group with a PCL-R score represented a higher risk profile overall. Their mean HCR-20 score was higher (*U* = 8881.50, *p* = .01); they were less likely to suffer

from psychotic disorders ($\chi^2(2) = 6.40$, $p = .01$), but more likely to have personality disorders ($\chi^2(2) = 10.51$, $p < .01$), more specifically, Cluster B personality disorders ($\chi^2(2) = 4.42$, $p = .04$), and antisocial personality disorder ($\chi^2(2) = 15.42$, $p = .01$). They also had on average more convictions ($U = 29833.00$, $p = .01$) and they were younger at the time of their first conviction ($U = 30077.00$, $p = .01$). The subsample with multiple PCL-R scores for the most part did not differ from the subsample with only one PCL-R score, with the exception of antisocial personality diagnosis, which was more common in the subsample with multiple scores ($\chi^2(2) = 5.72$, $p = .02$).

Measures

PCL-R

The PCL-R (Hare, 2003; Dutch translation: Vertommen, Verheul, de Ruiter, & Hildebrand, 2002) has become the standard scale to assess the degree of psychopathy in forensic settings. The scale consists of 20 items, scored according to the degree of applicability (0 = *not present*, 1 = *possibly present*, 2 = *definitely present*), yielding scores ranging from 0 to 40. Higher scores indicate a higher level of psychopathic traits. Categorically, Hare (2003) recommended a cut-off of 30, although research in European countries has suggested that scores of 25 or above are indicative of psychopathy (Cooke & Michie, 1999). Although often used categorically, the dimensional nature of the construct assessed by PCL-R scores has been evident in numerous taxometric studies, indicating that particular diagnostic cut scores are somewhat arbitrary distinctions rather than indicative of theoretically important, naturally occurring subgroups of psychopaths and nonpsychopaths (Edens, Marcus, Lilienfeld, & Poythress, 2006; Guay, Ruscio, Knight, & Hare, 2007; Murrie et al., 2007; Walters, Duncan, & Mitchell-Perez, 2007; Walters, Marcus, Edens, Knight, & Sanford, 2011). Ideally, the PCL–R is scored using file information and a semistructured interview with the offender. In the current study, unfortunately, the scoring procedure (with or without interview) was not noted in the files. It is likely, however, that in most cases, the PCL-R was scored with an interview because this is the method mostly applied in the field, compared with research studies that more typically rely only on file data (Hawes et al., 2013). Based on the informal survey of MSU and prison staff conducted by the authors, typically an interview would be performed unless the patient refused to participate. In case of missing items, factor and facet scores were prorated as recommended in the manual (Hare, 2003). The mean score ($M = 21.1$, *SD*

= 6.6) in the present sample was comparable to the forensic psychiatric population described in the PCL-R manual (*M* = 21.5, *SD* = 6.9).

Recidivism

Data concerning new criminal sentences were obtained from the Central Criminal Records of the Ministry of Justice. Because recidivism data solely based on new sentences would give an incomplete base rate of offending in this particular population (e.g., 8.9% violent and 15.2% general recidivism), we also included previously described crime-related incidents (see description under "Setting") that were reported directly to the CPS in order to capture a more accurate and comprehensive record of subsequent criminal activities. Crime-related incidents were defined as incidents coded under offending categories, whether or not they ultimately lead to further prosecution or sentencing. The prosecutor decides on whether the charges will be prosecuted. It is possible that because the person already is interned for an indefinite period, new offenses will not be prosecuted. As a result, using only the official recidivism data would have resulted in an artificially low base rate. Likewise, it is plausible that a sentencing date occurred without being previously reported as an incident. Recidivism is therefore coded yes if an incident was reported to the CPS, if the person recidivated, or both. To be consistent with other recidivism research, this combined outcome measure was divided into general and violent recidivism (including sexual recidivism). The base rates in the study group for violent and general recidivism were 42.9% and 62.9%, respectively. Follow-up times for the whole group ranged from 368 days to 3609 days (*M* = 1950.9 days, *SD* = 878.6 days). For prison scores follow-up times ranged from 385 days to 3,609 days (*M* = 1874.6, *SD* = 893.7 days) and for hospital scores from 368 days to 3,424 days (*M* = 1911.2, *SD* = 871.3 days). To correct for these large time differences in follow-up periods predictive validity was analyzed using a fixed follow-up period of 2 years. This reduced the sample to 203 participants, with 23.2% violent and 36.9% general recidivism.

Procedure

Data regarding level of education, psychiatric diagnosis, criminal history, hospitalization/imprisonment periods, risk assessment scores, and IQ scores were gathered by accessing both CPS files and psychiatric hospital records. Diagnoses were based on the

Diagnostic and Statistical Manual of mental disorders-IV-Text Revision (DSM–IV–TR; American Psychiatric Association, 2000). The PCL-R scores found in the files were field scores rated by prison psychologists or psychiatrists while the participants stayed in prison either during pretrial assessment or as part of the conditional release process (which we referred to as *prison scores*), and/or during a stay in a forensic psychiatric hospital by criminologists (criminal justice professionals with a master's degree) or psychologists (which we referred to as *hospital scores*). In other words, the PCL-R was scored as part of routine forensic clinical practice, prior in time to the outcome measures, which were later on collected by the researchers. Unfortunately, additional information on individual raters such as age, sex, or years of experience was not available.

In 176 cases, only one PCL-R score was available and in 48 cases, multiple PCL-R scores were available. For the analyses comprising repeated measures, all double scores were used: 38 participants had two scores (n = 38 total comparisons), eight participants had three scores resulting in three possible comparisons (n = 24 total comparisons), and two cases with four scores resulting in six possible comparisons (n = 12 total comparisons), adding up to a total of 74 pairs (see Figure 1). Among these pairs, 34 consisted of scores from the same type of institution (e.g., PCL-R_1 hospital vs. PCL-R_2 hospital or PCL-R_1 prison vs. PCL-R_2 prison) and 38 came from different institutions (e.g., PCL-R_1 hospital vs. PCL-R_2 prison or PCL-R_1 prison vs. PCL-R_2 hospital).

To examine predictive validity, the first PCL-R score of each participant was included in the analyses. In addition, when looking at the PCL-R scores stratified across settings, the first score within settings (prison scores, n = 69 vs. hospital scores, n = 180) was included in the analyses.

As in most archival file studies, the dataset contained some missing values. Factor scores and especially facet scores were not always reported along with the PCL-R total scores, resulting in lower samples sizes for these analyses, as noted in the tables.

Ethical approval was obtained from the Ethics committee of Antwerp University Hospital.

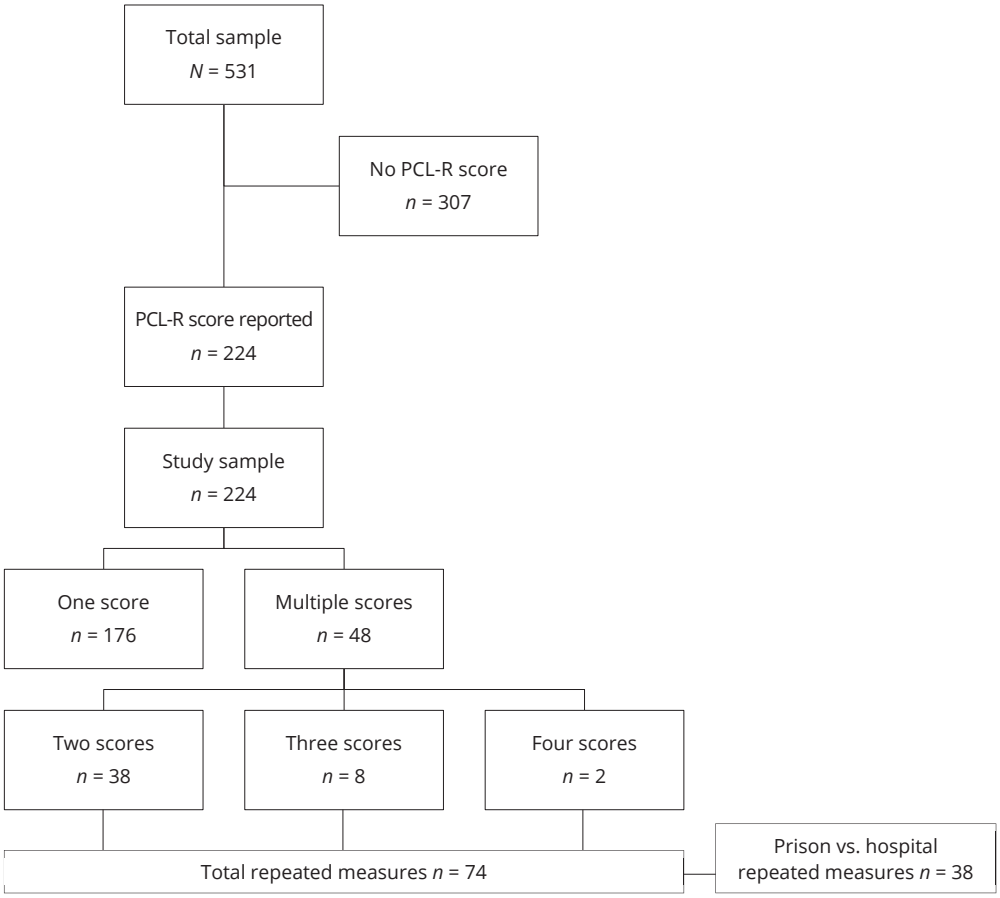


Figure 1. Availability of Psychopathy Checklist-Revised (PCL-R scores.

Statistical analyses

Unless otherwise noted, SPSS version 22 (IBM Corp., 2013) was used for the statistical analyses. Missing values were handled using pairwise deletion. Subsamples were compared using chisquare or Fisher Exact for nominal variables and Mann–Whitney for non-normally distributed continuous data. Independent samples t-tests and paired samples t-tests were used for the comparison of normally distributed data.

Absolute difference scores were computed by subtracting the repeated PCL-R scores from each other. An IRR analysis using the Kappa statistic was conducted to determine consistency among categorical ratings. The guidelines in Landis and Koch (1977) were used to interpret the different Kappa scores: < 0 poor, 0–0.2 slight; .21–.40 fair; .41–.60 moderate; .61–.80 substantial; and .81–1.00 (almost) perfect agreement. Two-way random intraclass correlation coefficients ($ICC_{A,1}$ absolute agreement, single measurement) were used to evaluate the IRR for PCL-R total, factor and facet scores. Two-way random intraclass correlation models both an effect of rater and of score (i.e., two effects) and the model assumes both are drawn randomly from larger populations. Absolute agreement ICCs are important for the PCL-R because total scores have a well-established clinical meaning. Although absolute agreement ICCs consider differences in the actual scores to be errors, consistency agreement ICCs focus on whether the evaluators tended to give the highest or lowest scores to the same offenders. In general, ICCs tend to be higher for consistency agreement than absolute agreement and higher for scores averaged across multiple evaluators than a single evaluator. Fleiss (1986) critical values for single measures were used: ICC < .40 = poor; ICC ≥ .40 = moderate; ICC ≥ .60 = good; and ICC ≥ .75 = excellent.

An instrument's standard error of measurement (SEM) can be used as an estimate of "the reasonable limits of the true score" (Anastasi, 1982). The rationale of SEM is that if 100 hypothetical raters assessed the same participant at the same time, about 68% of those scores would fall within one SEM unit of the participant's obtained actual score and about 95% would fall within two SEMs (Anastasi, 1982). The SEM used for the current study comes from the Hare (2003) manual, which reported a SEM of 2.7 for male forensic psychiatric patients. The percentage of scores falling within one and two SEMs was calculated to determine whether these percentages corresponded to what would be predicted based on the reliability statistics reported in the Hare (2003) manual. To examine predictive validity, the choice was made to use ROC analyses (instead of e.g., survival analyses) in order to facilitate comparison with previous studies. Differences in predictive validity between settings (hospital vs. prison) were assessed using the ROCOMP command in STATA 12 (StataCorp, 2011). Interpretation of the magnitude of the effect size of the corresponding AUC-values was determined using the classification of Rice and Harris (2005): AUC ≥.56 = little effect; AUC ≥ .64 = moderate effect; and AUC ≥.71 = large effect.

RESULTS

Descriptive statistics

Mean PCL-R total, factor and facet scores for the entire sample and the different settings are reported in Table 1. In 22.3% (*n* = 50/224) of the cases, no or only partial information on item

Table 1. Mean PCL-R scores for the entire population and stratified per setting for all scores

	Setting	*n*	Mean (*SD*)	Range	*t*	*p*	*d* [CI]
Total							
	All scores	224	21.1 (6.6)	6–36.8			
	Hospital scores	180	21.4 (6.2)	6–36.8			
	Prison scores	69	21.4 (7.3)	9–35.8	0.02	.98	.00 [-.27, .28]
Factor 1							
	All scores	215	8.6 (3.4)	1.1–16			
	Hospital scores	180	8.7 (3.3)	2–16			
	Prison scores	61	8.5 (3.6)	1.1–16	0.38	.70	.05 [-.23, .35]
Factor 2							
	All scores	213	10.3 (3.9)	1–18			
	Hospital scores	179	10.5 (3.7)	1–18			
	Prison scores	60	10.3 (4.0)	1–18	0.41	.68	005 [-.23, .35]
Facet 1							
	All scores	172	3.1 (2.1)	0–8			
	Hospital scores	157	3.2 (2.1)	0–8			
	Prison scores	37	2.7 (2.2)	0–8	1.31	.19	.19 [-.15, .57]
Facet 2							
	All scores	173	5.6 (1.9)	0–8			
	Hospital scores	157	5.7 (1.8)	0–8			
	Prison scores	38	5.5 (2.2)	1–8	0.39	.70	.11 [-.24, .47]
Facet 3							
	All scores	170	6.1 (2.5)	0–10			
	Hospital scores	154	6.1 (2.4)	0–10			
	Prison scores	38	6.1 (2.5)	1–10	0.13	.90	.02 [-.33, .38]
Facet 4							
	All scores	163	5.5 (2.6)	0–10			
	Hospital scores	150	5.7 (2.5)	0–10			
	Prison scores	34	4.8 (2.7)	0–10	1.95	.05	.29 [.27, .64]

CI = confidence interval
* $p < .05$.

scores was available. No significant difference was found regarding total PCL-R score for those cases with factor or facet scores reported and those without. In the cases where item scores were available (77.7%, *n* = 174/224), none of the items had more than 10% missing, except for Item 12 E*arly Behavior Problems* (19%, *n* = 33/174). On Item 12, a larger amount was missing for the prison scores (31.6%, *n* = 12/38) compared with the hospital scores (17.2%, *n* = 27/157). The mean PCL-R for the entire sample was 21.1 (*SD* = 6.6) with scores ranging from 6 to 36.8. The mean score for Factor 1 was 8.6 (*SD* = 3.4, range 1.1 –16) and for Factor 2 10.3 (*SD* = 3.9, range = 1–18). None of the scores differed significantly between settings. Almost one out of 10 participants (8.9%, *n* = 20/224) scored above the cut-off of 30 and one third (33.5%, *n* = 75/224) had a score of 25 or above on the total PCL-R score. Stratified per setting, the percentage above

Table 2. Mean PCL-R scores repeated measures stratified per setting

	Setting	*n*	Mean (*SD*)	Range	*t*	*p*	*d* [CI]
Total							
	Hospital	38	22.9 (5.8)	8.2–34			
	Prison	38	20.6 (6.8)	11–34	–2.04	.049*	.33 [-.13, .78]
Factor 1							
	Hospital	35	9.2 (2.9)	3–15			
	Prison	35	8.6 (3.5)	3–16	–1.11	.28	.19 [-28, .66]
Factor 2							
	Hospital	34	10.3 (3.7)	1–16.7			
	Prison	34	9.5 (3.7)	2–16	–1.34	.19	.23 [-.26, .69]
Facet 1							
	Hospital	23	2.9 (2.1)	0–8			
	Prison	23	3.2 (2.1)	0–8	1.10	.28	.32 [-.26, .91]
Facet 2							
	Hospital	23	5.9 (1.7)	2–8			
	Prison	23	5.1 (1.9)	2–8	–1.68	.11	.35 [-.20, .97]
Facet 3							
	Hospital	23	5.3 (2.1)	1–8.7			
	Prison	23	5.3 (2.6)	1–10	–0.02	.98	.00 [-.58, .57]
Facet 4							
	Hospital	19	5.2 (3.12)	0–10			
	Prison	19	4.1 (2.4)	0–9	–1.98	.06	45 [-.19, 1.10]

* $p < .05$.

the 30 cut-off was more than double in prison settings, compared with hospital settings (15.9%, *n* = 11/69 vs. 7.2%, *n* = 13/180, *p* = .04), whereas the percentage above the 25 cut-off did not differ significantly (36.2%, *n* = 25/69 for prison scores vs. 33.9%, *n* = 61/180 for hospital scores, *p* = .73). When repeated measures mean scores for the same offender were compared across settings (see Table 2), no significant differences were found for the mean scores except for the total PCL-R score, which was marginally higher for hospital scores (*M* = 22.9, *SD* = 5.8) compared with prison scores (*M* = 20.6, *SD* = 6.8), *t*(37) =-2.04, *p* = .049).

Interrater reliability analyses

Repeated measures ICC

Mean difference scores of the PCL-R total, factor and facet scores for the repeated measures can be found in Table 3, together with the ICCs. Analyses were repeated stratified per setting: scores from the same setting (hospital vs. hospital or prison vs. prison) and scores from different settings (hospital vs. prison). Difference scores for all repeated measures ranged from 0 to 19 for the total score, with an average of 5.8 (*SD* = 4.3). Stratified per setting, the mean difference score for repeated measures within settings was (*M* = 5.8, *SD* = 4.3) and for the scores from different settings (*M* = 5.8, *SD* = 4.4). For the PCL-R total score 28.4% (*n* = 28/74) of the scores fell within 1 SEM, 54.1% (*n* = 40/74) within 2 SEMs and 45.9% (*n* = 34/74) fell more than 2 SEMs apart. For all repeated measures, more than a 10-point difference on the total score was found in 16.2% (*n* = 12/74) of the cases. When looking at the scores stratified per setting, 29.4% (*n* = 10/34) of the scores from the same setting fell within 1 SEM, 52.9% (*n* = 18/34) within 2 SEMs, and 47.1% (*n* = 16/34) fell above 2 SEMs. Comparing scores from different settings, 26.3% (*n* = 10/38) of the scores fell within 1 SEM, 55.3% (*n* = 21/38) within 2 SEMs, and 44.7% (*n* = 17/38) fell above 2 SEMs. More than a 10-point difference on the total score was found across settings in 13.2% (*n* = 5/38) of the cases and within settings in 17.6% (*n* = 6/34) of the cases. The absolute agreement ICC for a single PCL-R total rating was .42 (95% CI [.22, .59]), and .44 (95% CI [.22, .62]) for both Factor 1 and Factor 2. Stratified per setting, the $ICC_{A,1}$ for same setting total PCL-R scores was .47 (95% CI [.17, .70]) and for different setting scores .37 (95% CI [.06, .61]). In same settings the $ICC_{A,1}$ for Factor 1 (.39, 95% CI [.04, .66]) was higher than for Factor 2 (.28, 95% CI [-.06, .57]). The opposite pattern was found in different settings where the $ICC_{A,1}$ for Factor 2 (.57, 95% CI [.30, .76]) was higher

than for Factor 1 (.50, 95% CI [.20, .71]). The ICCs at item level are presented in Table 4. Using the interpretive guidelines suggested by Fleiss (1986), most of the ICCs (60%, *n* = 12/20) were in the poor range, 35% (*n* = 7/20) were in the moderate range and only one $ICC_{A,1}$ was good (Item 18). Similar ICCs were found when IRR was calculated controlling for dependence of observations using a multilevel random regression model (results available upon request).

Table 3. Mean difference scores repeated measures and ICCs

	Setting	*n*	Mean difference (*SD*)	Range	ICC	CI
Total						
	All pairs	74	5.8 (4.3)	0–19	.42	.22–.59
	Same setting	34	5.8 (4.3)	0–16.7	.47	.17–.70
	Different setting	38	5.8 (4.4)	0–19	.37	.06–.61
Factor 1						
	All pairs	65	3.0 (2.0)	0–7	.44	.22–.62
	Same setting	30	3.2 (2.4)	0–7	.39	.04–.66
	Different setting	35	2.8 (1.7)	0–7	.50	.20–.71
Factor 2						
	All pairs	64	3.3 (2.1)	0–9.1	.44	.22–.62
	Same setting	30	3.9 (2.2)	0.8–9.1	.28	-.06–.57
	Different setting	34	2.9 (2.0)	0–7	.57	.30–.76
Facet 1						
	All pairs	40	1.6 (1.3)	0–5	.54	.28–.73
	Same setting	17	2.0 (1.5)	0–5	.33	-.19–.69
	Different setting	23	1.2 (1.1)	0–5	.70	.41–.86
Facet 2						
	All pairs	40	1.6 (1.5)	0–5	.27	-.03–.53
	Same setting	17	1.7 (1.5)	0–4	.24	-.28–.65
	Different setting	23	1.6 (1.5)	0–5	.28	-.12–.61
Facet 3						
	All pairs	40	2.2 (1.7)	0–7	.28	-.04–.55
	Same setting	17	2.0 (1.6)	0–7	.17	-.36–.60
	Different setting	23	2.3 (1.8)	0–6	.24	-.21–.59
Facet 4						
	All pairs	29	2.0 (1.3)	0–6	.60	.30–.79
	Same setting	10	1.6 (0.9)	0–3	.65	.07–.90
	Different setting	19	2.2 (1.5)	0–6	.59	.20–.82

Repeated measures kappa

The IRR, as measured by the Kappa statistic (*n* = 74), for categorical data using a cut-off of 30 was .22 (*p* = .06) and .13 when using a cut-off of 25 (*p* = .25). Stratified per setting, there was less agreement on the categorized individuals with high psychopathic traits based on a cut-off of 30 when assessed in the same setting (*n* = 34) (Kappa = .11, *p* = .54) compared with different settings (*n* = 38; kappa= .31, *p* = .047). Stratified per setting, poor agreement on categorizing individuals with high psychopathic traits based on a cut-off of 25 in the same setting (*n* = 34) (Kappa = .19, *p* = .26) was comparable with different settings (*n* = 38) (Kappa = .07, *p* = .69). Using the interpretive guidelines suggested by Landis and Koch (1977), agreement was slight for a cut-off of 25 and slight to fair for a cut-off of 30.

Table 4. Item ICCs repeated measures

Item	*N*	$ICC_{A,1}$	Fleiss criteria
Item 18	30	.66	good
Item 17	30	.56	moderate
Item 19	40	.53	moderate
Item 6	41	.52	moderate
Item 2	41	.49	moderate
Item 4	40	.49	moderate
Item 11	35	.46	moderate
Item 12	24	.42	moderate
Item 7	41	.38	poor
Item 10	38	.37	poor
Item 3	41	.36	poor
Item 14	41	.33	poor
Item 9	38	.30	poor
Item 1	41	.27	poor
Item 5	41	.26	poor
Item 8	41	.21	poor
Item 16	41	.20	poor
Item 15	40	.19	poor
Item 13	40	.17	poor
Item 20	40	.14	poor

Predictive validity analyses

Predictive validity was assessed for all scores and for the repeated measures. First, Table 5 reports AUC values for PCL-R total, factor and facet scores of all scores after a fixed follow-up period of 2 years. Similar (although somewhat lower) AUCs were found when repeating the

Table 5. Area under the curve for the entire population and stratified per setting for all scores using a fixed follow-up period of two years

			General recidivism				Violent recidivism			
	Setting	*n*	AUC	*p*	CI	*p**	AUC	*p*	CI	*p**
Total										
	All	203	.59	.10	.49–.68		.57	.14	.48–.66	
	Hospital	162	.58	.10	.49–.67		.59	.10	.49–.69	
	Prison	63	.63	.08	.49–.78	.52	.55	.57	.39–.70	.62
Factor 1										
	All	195	.51	.90	.40–.61		.51	.79	.42–.61	
	Hospital	162	.49	.78	.39–.58		.53	.58	.43–.64	
	Prison	56	.61	.15	.46–.76	.16	.52	.82	.36–.68	.91
Factor 2										
	All	194	.63	.02	.53–.72		.62	.01	.54–.71	
	Hospital	162	.64	.00	.55–.72		.59	.10	.50–.69	
	Prison	55	.69	.02	.55–.84	.49	.62	.17	.46–.77	.79
Facet 1										
	All	155	.50	.95	.38–.63		.45	.38	.34–.56	
	Hospital	139	.47	.49	.37–.57		.49	.86	.37–.61	
	Prison	34	.47	.81	.26–.69	.94	.33	.14	.11–.55	.21
Facet 2										
	All	156	.58	.21	.46–.70		.57	.21	.47–.68	
	Hospital	139	.50	.92	.40–.59		.55	.41	.44–.66	
	Prison	35	.69	.06	.52–.87	.05	.69	.11	.49–.89	.24
Facet 3										
	All	153	.63	.04	.51–.75		.63	.03	.53–.74	
	Hospital	136	.62	.02	.52–.71		.62	.06	.51–.73	
	Prison	35	.63	.22	.42–.84	.93	.53	.78	.28–.79	.57
Facet 4										
	All	147	.59	.14	.49–.69		.56	.28	.47–.66	
	Hospital	133	.59	.07	.50–.69		.52	.77	.42–.62	
	Prison	31	.65	.18	.44–.86	.65	.64	.24	.44–.84	.28

*Note. p** = significance value associated with the difference between prison and hospital AUCs.

analyses for the entire population with only a 1-year follow-up period (results available upon request). For general recidivism, the Factor 2 scores for the entire population were significant (AUC = .63, p = .02), as were the prison (AUC = .69, p = .02) and hospital (AUC = .64, p = .001) Factor 2 scores. Factor 2 scores also predicted violent recidivism in the entire population (AUC = .62, p = .01). At the facet level, the only predictive facets for general recidivism were the lifestyle facet (Facet 3) for the entire population (AUC = .63, p = .04), as well as the hospital scores (AUC = .62, p = .02). The lifestyle facet (Facet 3) also predicted violent recidivism for the entire population (AUC = .63, p = .03). Despite differences in predictive validity between settings, when testing AUC values between settings, these differences did not reach statistical significance.

Next, repeated measures across settings were compared in order to assess whether predictive validity would differ. When only multiple total and factor scores for the same offender were compared across settings (see Table 6), only the hospital total and Facet 3 scores were predictive of general recidivism. Again, differences found between settings did not reach statistical significance.

DISCUSSION

This is the first study to examine the field reliability and validity of PCL-R scores in Belgium, using clinical scores that were retrieved from archival data. The study further aimed to explore whether contextual pressure would occur when comparing scores obtained within a prison setting compared with a hospital setting and whether this would result in different predictive validity across settings. The general conclusion of the current study in this sample of forensic psychiatric patients was that the PCL-R in real world settings conducted by real world raters in Belgium is fairly unreliable although there was some evidence of modest to moderate predictive validity for Factor 2 scores. These results are generally consistent with a growing body of field research suggesting that the high levels of reliability reported in many controlled research studies are not generalizable to practice settings.

Table 6. Area under the curve repeated measures using a fixed follow-up period of two years

			General recidivism				Violent recidivism			
	Setting	*n*	AUC	*p*	CI	*p**	AUC	*p*	CI	*p**
Total										
	Hospital	34	.71	.04	.53–.89		.66	.14	.44–.88	
	Prison	34	.49	.92	.28–.70	.13	.35	.19	.15–.56	.05
Factor 1										
	Hospital	31	.60	.35	.39–.81		.51	.91	.27–.76	
	Prison	31	.57	.54	.34–.79	.82	.36	.26	.14–.60	.39
Factor 2										
	Hospital	30	.65	.16	.45–.86		.61	.37	.38–.84	
	Prison	30	.59	.42	.37–.81	.68	.54	.74	.31–.78	.70
Facet 1										
	Hospital	20	.58	.59	.28–.88		.45	.78	.08–.83	
	Prison	20	.43	.62	.15–.71	.50	.22	.09	.00–.44	.28
Facet 2										
	Hospital	20	.60	.48	.35–.86		.69	.26	.46–.91	
	Prison	20	.56	.68	.30–.82	.82	.55	.78	.26–.83	.43
Facet 3										
	Hospital	20	.80	.04	.52–1.00		.68	.28	.31–1.00	
	Prison	20	.55	.74	.23–.86	.27	.48	.89	.09–.86	.51
Facet 4										
	Hospital	16	.63	.43	.32–.93		.74	.16	.46–1.00	
	Prison	16	.65	.37	.31–.98	.94	.82	.06	.62–1.00	.65

Note. p^* = significance value associated with the difference between prison and hospital AUCs.

Descriptive statistics

The average total PCL-R score found in all patients with a PCL-R score was similar to the mean score for forensic psychiatric patients reported by Hare (2003), and there was not a significant difference found in patients scored within a prison versus a hospital setting (see Table 1). However, when comparing repeated measures for the same offender across settings (see Table 2), mean prison scores were lower than mean hospital scores, suggesting contextual pressure as expected. At the time of this study, high security forensic psychiatric beds were not available and the Belgian government received several convictions by the European Court of Human Rights for the practice of keeping internees in prison without appropriate psychiatric treatment. Therefore,

it was anticipated that prison raters would assign a lower score in order to facilitate transfer to a MSU.

Interrater reliability

When the repeated measures were further explored both across and within settings, substantial differences were found, with only a minority of the difference scores falling within 1 SEM. Although differences of more than six points – or two SEMs – should be extremely rare, this was the case in almost half of the scores. Our findings tended to fall between other recent field validity studies that have reported relatively high percentages of cases with differences exceeding two SEMs (Edens et al., 2015; Murrie et al., 2008; Sturup et al., 2014). In addition, mean differences in PCL-R scores were generally comparable with those reported in other field validity studies (DeMatteo, Edens, Galloway, Cox, Smith, & Formon, 2014; Edens et al., 2015; Miller et al., 2012; Murrie et al., 2008; Sturup et al., 2014). The ICCs reported in the current study ranged from .24 to .70 and were approximately half of those reported by Hare (2003) (e.g., ICC total PCL-R score male forensic patients = .88 vs. .42). They were, however, in line with other field validity studies in North American samples (Boccaccini et al., 2014; Boccaccini et al., 2008; DeMatteo, Edens, Galloway, Cox, Smith, & Formon, 2014; Edens et al., 2015), although lower in comparison with another recent European field study (Sturup et al., 2014) reporting a total PCL-R score $ICC_{A,1}$ of .70. The higher scores found in Sweden could be explained by the fact that the PCL-Rs were administered by the same, independent government agency using a team approach and thereby minimizing partisanship effects that could influence the PCL-R score.

Following Fleiss's (1986) criteria for interpreting reliability statistics, only Facet 4 ($ICC_{A,1}$ = .59 –.65) would be considered "good" which was in line with other field validity studies (Miller et al., 2012; Sturup et al., 2014), which found the highest ICCs in Facet 4 scores. However, according to suggested interpretive guidelines for applied settings, none of the ICCs reached an acceptable level (Heilbrun, 1992; Nunnally, 1978; Rosenthal & Rosnow, 1991). Facet 2 and 3 showed particularly low IRR ($ICC_{A,1}$ = .17–.28). Surprisingly, Factor 2 scores obtained within the same institution had lower IRR compared with scores obtained from different institutions ($ICC_{A,1}$ = .28 vs. .57). Contrary to other field validity studies finding more rater agreement for Factor 2 scores as compared with Factor 1 scores, ICCs for factor scores showed mixed findings in the current

study. Comparing different settings, ICCs for Factor 2 scores were higher than Factor 1, while the opposite pattern emerged within same setting comparisons. Overall there was no difference. As was the case in other field validity studies, the highest item-level ICC was found for Item 18. Surprisingly and contrary to other field validity studies, the lowest ICC was found for Item 20 (Miller et al., 2012; Sturup et al., 2014).

Another way to examine rater agreement is at the categorical level. Among the eight cases with one score at or above 30, only three (37.5%) were above this cut-off on the second assessment. Using a cut-off of 25, in only 11 out of 26 cases (42.3%) did both raters score the examinee as at or above a score of 25. The kappa coefficients were in the poor range. Other studies (DeMatteo, Edens, Galloway, Cox, Smith, & Formon, 2014; Sturup et al., 2014) have also found lower kappa values than reported in the Hare manual (Kappa = .67, Hare, 2003) but still higher than in the current study. Stratified per setting, there was less agreement on categorizing individuals with high psychopathic traits based on a cut-off of 30 in the same setting (Kappa = .11, p = .54) compared with different settings (Kappa = .31, p = .047). However, when using the cut-off of 25 the differences between settings disappeared.

In sum, rater agreement was poor. Numerous possible explanations exist for these findings. First, our sample consisted mainly of nonsexual forensic psychiatric offenders, which may have played a role, because Edens et al. (2015) found that ICCs in their nonsexual offender subsample ($ICC_{A,1}$ = .46) was lower compared with their sexual offender subsample ($ICC_{A,1}$ = .66). Also, it is possible that our sample comprised more complex cases, because a PCL-R was not scored on a routine basis. Another explanation relates to rater differences. The small number of completed PCL-R scores found over an extensive study period lead us to conclude that there were few prolific raters in our sample. There are some differences between the two settings that could further explain the discrepancies in results. In the hospitals the PCL-Rs were generally administered by criminologists (sometimes psychologists) and although all of them have followed a formal PCL-R training, they received the training at different institutes (e.g., some in Belgium, some in the Netherlands). Furthermore, because they are part of the multidisciplinary team, hospital evaluators are up-to-date on the daily behavior and progress of the patient. As a result their knowledge about the functioning of the patient at the moment of stay at the MSU as well as their therapeutic relationship with the patient could influence their scoring. In prison, PCL-R

interviews are generally administered by psychologists (sometimes psychiatrists) and they have received more uniform training. Our study is the only field research examining raters other than psychologists or psychiatrists, namely criminologists. Although these raters have a master's degree and an advanced understanding of crime, law, and criminal justice, as well as basic education in psychology and sociology, it is reasonable to assume that their knowledge on (diagnosing) personality disorders and mental health symptoms will be less extensive. Unfortunately, we could not link individual scores to the profession of the rater. Further research should focus on this matter.

Predictive validity

Although poor interrater reliability does not necessarily indicate that predictive validity will be low, this is generally what was found. When comparing AUC for the repeated measures across settings (Table 6), only the total and Facet 3 hospital score significantly predicted general recidivism, whereas none of the scores reached significance for violent recidivism. No significant difference in AUC was found, although the difference between the prediction of violent recidivism nearly reached significance in favor of the hospital scores (AUC = .66 vs. .35, p = .05). It should be noted however that these analyses were based on small numbers (Table 6). At this point no firm conclusions can be drawn from the current study regarding the superiority of different groups of raters in terms of predictive validity.

In line with the poor predictive validity of the repeated measures, a similar pattern was found when assessing the entire sample (Table 5). Overall, the predictive validity was poor, especially for total PCL-R score and Factor 1, which did not predict general or violent recidivism. Factor 2 scores significantly predicted general recidivism for all groups, whereas Factor 2 scores predicted violence only for the combined population (prison and hospital scores). On the facet level, surprisingly, Facet 3 scores were the only significant predictors of general (all and hospital scores) and violent recidivism (all scores). Although some AUCs reached statistical significance, the level ranged from small to moderate effect sizes, with moderate effect sizes in Factor 2 prison and hospital scores for predicting general recidivism and small effect sizes for all scores (Factor 2 and Facet 3) for predicting general and violent recidivism and for Facet 3 hospital scores for general recidivism. These small effects were somewhat surprising, given the extant research

demonstrating good predictive validity of the PCL-R in well controlled research designs with regard to general and violent recidivism (Hemphill, Hare, & Wong, 1998; Leistico et al., 2008; Salekin, Rogers, & Sewell, 1996; Walters, 2003; Yang et al., 2010) although results for Factor 1 in recent meta-analyses have been quite poor.

Although the lack of significant findings for prison AUCs could be explained by the small sample in the prison settings, the numbers for the hospital and total group were adequately powered. Also, our study results were in line with prior research identifying null findings with respect to the predictive validity in field settings (Murrie et al., 2012; Neal et al., 2015). So, the question remains why this is the case. Potential explanations relate to reliability issues raised earlier such as negligence or insufficient experience in scoring. Another potential explanation of the poor predictive validity is that the assessment, and thus the knowledge of a high psychopathy score and hence high risk status, could have resulted in more effective risk management, which would obviously attenuate effect sizes.

Clinical implications

Field validity studies such as the current one are important for researchers to consider when developing and refining new instruments and for clinicians to be aware of when conducting assessments in practice (Neal et al., 2015).

When discussing scores, raters and judges should be aware of the fact that potential biases of the rater could have an important impact. As was recently shown by Boccaccini, Chevalier, Murrie, and Varela (2015), when surveying evaluators who conducted SVP evaluations, this bias extends to score interpretation. They found considerable variability in the cut scores evaluators used to classify offenders as low, moderate, and high risk. For example, the cut scores for high psychopathy ranged from 20 to 37. Also, some of the variability in cut score use was explained by adversarial affiliation, with defense evaluators setting a higher threshold for high or moderate psychopathy compared to prosecution evaluators. Interestingly, raters acknowledge bias in general, whereas they are convinced that scores they assigned themselves were not influenced (Boccaccini et al., 2015). Unfortunately, in the current study no data were available regarding the person who administered the PCL-R (only place of administration) and therefore no analyses could be conducted to investigate issues such as the effects of level and type of training,

experience, prolific raters, method of assessment (file vs. interview), or personality traits of the evaluators. Our findings suggested that contextual pressures may play a small role (i.e., giving a more favorable or lower score in order to facilitate transfer from prison to a forensic hospital), but warrant further research because they were based on small numbers. We could not control for other variables but missing information on several potentially important issues did raise the question as to which raters are best fit for the job? Consideration of possible causes for low rater agreement identified in the literature review suggest a number of recommendations in terms of how one might improve clinical practice regarding the use of the PCL-R. First, perhaps raters who do not frequently use the PCL-R should not use it at all, because prolific raters tend to outperform scores from less prolific raters (Murrie et al., 2012). It might be advisable to have few evaluators doing all the ratings in settings where the PCL-R is not used on a regular basis. The question whether these evaluators should be psychologists, psychiatrists, or criminologists with or without a doctoral degree remains for the moment unanswered. Another question relates to the usage of an interview. Harris et al. (2013) suggested that clinicians conducting their own interviews partly explained why field reliability yielded poor results, because they are more prone to manipulation. The authors went on by stating that "the literature is clear—when the PCL-R is scored for violence risk assessment, at least, there would be no cost, and potentially some real benefit, to mandating scoring without interview" (p. 1359). We think that this is an important point that should be assessed in future field validity studies.

Measurement error is inherent to all psychometric instruments, but it should be kept to an absolute minimum. Given the present results and those from previous field studies it increasingly seems advisable that PCL-R interviews should be performed by preferably two independent, trained, and experienced raters, ideally by using a consensus score. These recommendations were already stated in the PCL-R manual (Hare, 2003), and have been repeated many times. For example, as Gacono and Hutton (1994) pointed out, reliable and valid ratings after formal PCL-R training are best achieved with a minimum of 10 further conjoint ratings, frequent use of the instrument and using two raters whenever possible. Furthermore, raters are encouraged to keep up with relevant liter- ature on the PCL-R and ignore introspection about etiology because for most criteria this would be distracting and interfering with obtaining a valid rating (Gacono & Hutton, 1994). In addition, after following initial training, continuous peer review processes can

minimize drift from item descriptions. In practice, however, it is often not possible (due to financial or time restraints) to adhere to all of the recommendations, as evidenced by the fact that most criminal cases seem to involve only one examiner's assessment results (DeMatteo, Edens, Galloway, Cox, Smith, Koller, et al., 2014; Edens et al., 2015). Nevertheless, raters should be aware that the utility of their PCL-R score (and the therapeutic and judicial decisions that stem from that score) will drop drastically when not following these recommendations.

Finally, with regard to the reporting of predictive validity findings, our results showed that the total PCL-R score was not especially useful, whereas Factor 2 and Facet 3 scores were. As such, reporting only a total PCL-R score paints only part of the picture. However, as noted by Boccaccini et al. (2015), most evaluators used categorical interpretations and few evaluators reported factor or facet scores. Forensic examiners should provide a comprehensive report of their PCL-R findings, including a discussion on the cut-off used and the profile of the facet scores. Depending on the assessment context, they might also consider not reporting Factor 1 scores at all unless there is some compelling reason for their inclusion.

Methodological considerations

Several methodological issues have to be taken into account when interpreting our findings. First, among only 42.2% (*n* = 224) of the forensic patients was a PCL-R score available and in only one third of these cases (21.4%, *n* = 48) were multiple scores present. A second possible bias relates to the time difference between the first and the second PCL-R score in the current study. For the whole group, the mean difference was 3.2 years However, time between assessments was not correlated with differences scores (r = .07, p = .59). As was the case in some other field validity studies (e.g., Sturup et al., 2014), the present study conflated IRR with temporal instability (see Murrie et al., 2008 for a discussion on test–retest and field validity). Although psychopathy is sometimes argued to be immutable and PCL-R scores are supposed to be relatively unchanging assessments of personality traits over the life-course, test–retest values for the PCL-R have actually been quite modest in some large-scale research projects. For example, a relatively low 2-year test–retest reliability value (ICC = .60 for men and ICC = .65 for women) was reported by Rutherford, Cacciola, Alterman, McKay, and Cook (1999). For men, Factor 1 (ICC = .43) was significantly less reliable than Factor 2 (ICC = .60). For women, Factor 2 (ICC = .50) was significantly

less reliable than Factor 1 (ICC = .63). However, no dramatic changes should be expected because PCL-R scores should represent lifelong patterns and typical functioning. In addition, while there is cross-sectional evidence of marginal decreases in PCL-R total and Factor 2 scores with increasing age (Hare, 2003, p. 61), when comparing the mean scores on the second scored PCL-R, total, Factor 1, and Factor 2 scores were higher than the first score, although the difference did not reach statistical significance at the factor level. However, mean age at the second score was still well below 40 years.

A third limitation relates to the lack of specific information concerning the scoring method – with or without interview – and the raters performing the assessments. As the individual raters could not be traced based on the available data, certain statistical analyses (i.e., generalizability theory) could not be performed using the present data and potential influences could not be analyzed. Equally, limited information was available on why multiple scores were present in some cases. It is possible that these were more complex cases. Also, the number of raters and whether or not some raters were more prolific compared to others was unknown.

Further, it was not known whether or not the second rater had knowledge of the first PCL-R score. As such, we should acknowledge that our results may overestimate the degree to which examiners would independently arrive at similar PCL-R scores, given that we do not know whether knowledge of a previous score impacted the results of a second evaluation due anchoring bias (i.e., if a previous score was known one would be more likely to give a similar score). One might see this as a limitation. However, if this were the case higher ICCs would be expected, not lower ICCs. Therefore, despite the fact that there was a possibility that the second rater had knowledge of the first score, ICCs were still too low.

CONCLUSION

The results of this field validity study revealed interrater disagreement well beyond what would be expected based on reliability data reported in the PCL-R professional manual (Hare, 2003) and adds to the growing evidence of poor ICCs in field settings. While the PCL-R leaves some room for subjectivity in scoring, the level of discrepancies found in the current study should raise serious concerns, particularly when considered in conjunction with the poor predictive validity results for

the total score and Factor 1. Clinicians using the PCL-R are encouraged to use ethical guidance in forensic evaluations and further examine the role that contextual pressures may play in their evaluations.

FOOTNOTES

1 The revised version of the HCR-20 (HCR^{V3}; Douglas, Hart, Webster, & Belfrage, 2013) does not require a PCL-R. In the revised version of the VRAG, the PCL-R was replaced by Facet 4 of the PCL-R (VRAG-R; Rice, Harris, & Lang).

2 A fourth case with an "unusually low" (Murrie et al., 2013, p. 1894) PCL-R score did not demonstrate significant differences across the two conditions.

3 The prosecutor decides on whether the charges will be prosecuted. It is possible that because the person already is interned for an indefinite period, new offenses will not be prosecuted. As a result, using only the official recidivism data would have resulted in an artificially low base rate. Likewise, it is plausible that a sentencing date occurred without being previously reported as an incident.

4 On one occasion different scores came from different hospitals; otherwise multiple scores within-hospital setting came from the same hospital.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Anastasi, A. (1982). *Psychological testing*. London: Collier Macmillan.

Bersoff, D. N. (2002). Some contrarian concerns about law, psychology, and public policy. *Law and Human Behavior, 26*(5), 565-574.

Boccaccini, M. T., Chevalier, C. S., Murrie, D. C., & Varela, J. G. (2015). Psychopathy Checklist-Revised use and reporting practices in sexually violent predator evaluations. *Sexual Abuse*. doi: 10.1177/1079063215612443

Boccaccini, M. T., Murrie, D. C., Rufino, K. A., & Gardner, B. O. (2014). Evaluator differences in Psychopathy Checklist-Revised factor and facet scores. *Law and Human Behavior, 38*(4), 337-345. doi: 10.1037/lhb0000069

Boccaccini, M. T., Turner, D. B., & Murrie, D. C. (2008). Do some evaluators report consistently higher or lower PCL-R scores than others? Findings from a statewide sample of sexually violent predator evaluations. *Psychology, Public Policy, and Law, 14*(4), 262-283. doi: 10.1037/a0014523

Boccaccini, M. T., Turner, D. B., Murrie, D. C., & Rufino, K. A. (2012). Do PCL-R scores from state or defense experts best predict future misconduct among civilly committed sex offenders? *Law and Human Behavior, 36*(3), 159-169. doi: 10.1037/h0093949

Cooke, D. J., Hart, S. D., & Michie, C. (2004). Cross-national differences in the assessment of psychopathy: do they reflect variations in raters' perceptions of symptoms? *Psychological Assessment, 16*(3), 335-339. doi: 10.1037/1040-3590.16.3.335

Cooke, D. J., & Michie, C. (1999). Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology, 108*(1), 58–68. doi: 10.1037/0021-843X.108.1.58

Cooke, D. J., Michie, C., & Hart, S. (2006). Facets of clinical psychopathy: Toward a clearer measurement. In C. Patrick (Ed.), *Handbook of psychopathy* (pp. 91–106). New York, NY: Guilford Press.

Costa, P. T. J., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Declercq, F., Willemsen, J., Audenaert, K., & Verhaeghe, P. (2012). Psychopathy and predatory violence in homicide, violent, and sexual offences: Factor and facet relations. *Legal and Criminological Psychology, 17*(1), 59-74. doi: 10.1348/135532510X527722

DeMatteo, D., Edens, J. F., Galloway, M., Cox, J., Smith, S. T., & Formon, D. (2014). The role and reliability of the Psychopathy Checklist-Revised in U.S. sexually violent predator evaluations: A case law survey. *Law and Human Behavior, 38*(3), 248-255. doi: 10.1037/lhb0000059

DeMatteo, D., Edens, J. F., Galloway, M., Cox, J., Smith, S. T., Koller, J. P., & Bersoff, B. (2014). Investigating the role of the Psychopathy Checklist-Revised in United States case law. *Psychology, Public Policy, and Law, 20*(1), 96-107. doi: 10.1037/a0035452

Douglas, K. S., Hart, S. D., Webster, C., & Belfrage, H. (2013). *HCR-20 (Version 3): Assessing Risk for Violence*. Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

Edens, J. F., Boccaccini, M. T., & Johnson, D. W. (2010). Inter-rater reliability of the PCL-R total and factor scores among psychopathic sex offenders: Are personality features more prone to disagreement than behavioral features? *Behavioral Sciences & the Law, 28*(1), 106-119. doi: 10.1002/bsl.918

Edens, J. F., Colwell, L. H., Desforges, D. M., & Fernandez, K. (2005). The impact of mental health evidence on support for capital punishment: Are defendants labeled psychopathic considered more deserving of death? *Behavioral Sciences & the Law, 23*(5), 603-625. doi: 10.1002/bsl.660

Edens, J. F., Cox, J., Smith, S. T., DeMatteo, D., & Sorman, K. (2015). How reliable are Psychopathy Checklist-Revised scores in Canadian criminal trials? A case law review. *Psychological Assessment, 27*(2), 447-456. doi: 10.1037/pas0000048

Edens, J. F., Davis, K. M., Fernandez Smith, K., & Guy, L. S. (2013). No sympathy for the devil: Attributing psychopathic traits to capital murderers also predicts support for executing them. *Personality disorders, 4*(2), 175-181. doi: 10.1037/a0026442

Edens, J. F., Marcus, D. K., Lilienfeld, S. O., & Poythress, N. G., Jr. (2006). Psychopathic, not psychopath: Taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology, 115*(1), 131-144. doi: 10.1037/0021-843X.115.1.131

Edens, J. F., Petrila, J., & Kelley, S. E. (in press). Legal and ethical issues in the assessment and treatment of psychopathy. In C. J. Patrick (Ed.), *Handbook of Psychopathy (2nd Ed.)*. New York, NY: Guilford.

Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. New York, NY: John Wiley.

Gacono, C. B., & Hutton, H. E. (1994). Suggestions for the clinical and forensic use of the Hare Psychopathy Checklist-Revised (PCL-R). *International Journal of Law and Psychiatry, 17*(3), 303-317.

Grann, M., Langstrom, N., Tengstrom, A., & Stalenheim, E. G. (1998). Reliability of file-based retrospective ratings of psychopathy with the PCL-R. *Journal of Personality Assessment, 70*(3), 416-426. doi: 10.1207/s15327752jpa7003_2

Guay, J. P., Ruscio, J., Knight, R. A., & Hare, R. D. (2007). A taxometric analysis of the latent structure of psychopathy: Evidence for dimensionality. *Journal of Abnormal Psychology, 116*(4), 701-716. doi: 10.1037/0021-843X.116.4.701

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist*. Toronto, ON: Multi-Health Systems.

Hare, R. D., Clark, D., Grann, M., & Thornton, D. (2000). Psychopathy and the predictive validity of the PCL-R: an international perspective. *Behavioral Sciences & the Law, 18*(5), 623-645.

Harris, G. T., Rice, M. E., & Cormier, C. A. (2013). Research and clinical scoring of the Psychopathy Checklist can show good agreement. *Criminal Justice and Behavior*, *40*(11), 1349-1362. doi: 10.1177/0093854813492959

Harris, P. B., Boccaccini, M. T., & Murrie, D. C. (2015). Rater differences in psychopathy measure scoring and predictive validity. *Law and Human Behavior, 39*(4), 321-331. doi: 10.1037/lhb0000115

Hawes, S. W., Boccaccini, M. T., & Murrie, D. C. (2013). Psychopathy and the combination of psychopathy and sexual deviance as predictors of sexual recidivism: Meta-analytic findings using the Psychopathy Checklist-Revised. *Psychological Assessment, 25*(1), 233-243. doi: 10.1037/a0030391

Heilbrun, K. (1992). The role of psychological testing in forensic assessment. *Law and Human Behavior, 16*(3), 257-272.

Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology, 3*, 139–170. doi: 10.1111/j.2044-8333.1998.tb00355.x

IBM Corp. (Released 2013). *IBM SPSS Statistics for Windows, Version 22.0*. Armonk, NY: IBM Corp.

Ismail, G., & Looman, J. (2016). Field inter-rater reliability of the Psychopathy Checklist-Revised. *International Journal of Offender Therapy and Comparative Criminology*. Advance online publication. http://dx .doi.org/10.1177/0306624X16652452

Jeandarme, I., Habets, P., Oei, T. I., & Bogaerts, S. (2016). Reconviction and revocation rates in Flanders after medium security treatment. *International Journal of Law and Psychiatry, 47*, 45-52. doi: 10.1016/j.ijlp.2016.02.033

Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior, 28*(4), 471-489. doi: 10.1177/009385480102800405

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159-174.

Laurell, J., Belfrage, H., & Hellstrom, A. (2010). Facets on the Psychopathy Checklist Screening Version and instrumental violence in forensic psychiatric patients. *Criminal Behaviour and Mental Health, 20*(4), 285-294. doi: 10.1002/cbm.779

Laurell, J., & Daderman, A. M. (2007). Psychopathy (PCL-R) in a forensic psychiatric sample of homicide offenders: Some reliability issues. *International Journal of Law and Psychiatry, 30*(2), 127-135. doi: 10.1016/j.ijlp.2004.08.011

Leistico, A. M., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the hare measures of psychopathy to antisocial conduct. *Law and Human Behavior, 32*(1), 28-45. doi: 10.1007/s10979-007-9096-6

Levenson, J. S. (2004). Reliability of sexually violent predator civil commitment criteria in Florida. *Law and Human Behavior, 28*(4), 357-368.

Lloyd, C. D., Clark, H. J., & Forth, A. E. (2010). Psychopathy, expert testimony, and indeterminate sentences: Exploring the relationship between Psychopathy Checklist-Revised testimony and trial outcome in Canada. *Legal and Criminological Psychology, 15*(2), 323-339. doi: 10.1348/135532509X468432

Miller, A. K., Rufino, K. A., Boccaccini, M. T., Jackson, R. L., & Murrie, D. C. (2011). On individual differences in person perception: Raters' personality traits relate to their Psychopathy Checklist-Revised scoring tendencies. *Assessment, 18*(2), 253-260. doi: 10.1177/1073191111402460

Miller, C. S., Kimonis, E. R., Otto, R. K., Kline, S. M., & Wasserman, A. L. (2012). Reliability of risk assessment measures used in sexually violent predator proceedings. *Psychological Assessment, 24*(4), 944-953. doi: 10.1037/a0028411

Murrie, D. C., Boccaccini, M. T., B., T. D., Meeks, M., Woods, C., & Tussey, C. (2009). Rater (dis)agreement on risk assessment measures in sexually violent predator proceedings: Evidence of adversarial allegiance in forensic evaluation? *Psychology, Public Policy and Law, 15*(1), 19-53.

Murrie, D. C., Boccaccini, M. T., Caperton, J., & Rufino, K. (2012). Field validity of the Psychopathy Checklist-Revised in sex offender risk assessment. *Psychological Assessment, 24*(2), 524-529. doi: 10.1037/a0026015

Murrie, D. C., Boccaccini, M. T., Guarnera, L. A., & Rufino, K. A. (2013). Are forensic experts biased by the side that retained them? *Psychological Science, 24*(10), 1889-1897. doi: 10.1177/0956797613481812

Murrie, D. C., Boccaccini, M. T., Johnson, J. T., & Janke, C. (2008). Does interrater (dis)agreement on Psychopathy Checklist scores in sexually violent predator trials suggest partisan allegiance in forensic evaluations? *Law and Human Behavior, 32*(4), 352-362. doi: 10.1007/s10979-007-9097-5

Murrie, D. C., Marcus, D. K., Douglas, K. S., Lee, Z., Salekin, R. T., & Vincent, G. (2007). Youth with psychopathy features are not a discrete class: A taxometric analysis. *Journal of Child Psychology and Psychiatry, 48*(7), 714-723. doi: 10.1111/j.1469-7610.2007.01734.x

Neal, T., Miller, S. L., & Shealy, R. C. (2015). A field study of a comprehensive Violence Risk Assessment battery. *Criminal Justice and Bahavior, 42*(9), 952-968. doi: 10.1177/0093854815572252

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Porter, S., Woodworth, M., Earle, J., Drugge, J., & Boer, D. (2003). Characteristics of sexual homicides committed by psychopathic and nonpsychopathic offenders. *Law and Human Behavior, 27*(5), 459-470. doi: 10.1023/A:1025461421791

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's *d*, and *r*. *Law and Human Behavior, 29*(5), 615–620. doi: 10.1007/s10979-005-6832-7

Rice, M. E., Harris, G. T., & Lang, C. (2013). Validation of and revision to the VRAG and SORAG: The Violence Risk Appraisal Guide-Revised (VRAG-R). *Psychological Assessment*, *25*, 951–965. http://dx.doi.org/10.1037/a0032878

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York, NY: McGraw-Hill.

Rufino, K. A., Boccaccini, M. T., Hawes, S. W., & Murrie, D. C. (2012). When experts disagreed, who was correct? A comparison of PCL-R scores from independent raters and opposing forensic experts. *Law and Human Behavior*, *36*, 527–537. http://dx.doi.org/10.1037/h0093988

Rutherford, M., Cacciola, J. S., Alterman, A. I., McKay, J. R., & Cook, T. G. (1999). The 2-year test-retest reliability of the Psychopathy Checklist Revised in methadone patients. *Assessment, 6*(3), 285-292.

Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the Psychopathy Checklist and Psychopathy Checklist-Revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice, 3*(3), 203–215. doi: 10.1111/j.1468-2850.1996.tb00071.x

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499-513. doi: 10.1016/j.cpr.2010.11.009

StataCorp. (2011). *Stata Statistical Software: Release 12*. College Station, TX: StataCorp LP.

Sturup, J., Edens, J. F., Sorman, K., Karlberg, D., Fredriksson, B., & Kristiansson, M. (2014). Field reliability of the Psychopathy Checklist-Revised among life sentenced prisoners in Sweden. *Law and Human Behavior, 38*(4), 315-324. doi: 10.1037/lhb0000063

Sundby, S. E. (1997). Capital jury and absolution: The intersection of trial strategy remorse and the death penalty. *Cornell Law Review, 83*, 1557.

Vertommen, H., Verheul, R., Ruiter, C. de, & Hildebrand, M. (2002). *Handleiding bij de herziene versie van Hare's Psychopathie Checklist*. Lisse: Swets Test Publishers.

Vogel, V. de, Ruiter, C. de, Hildebrand, M., Bos, B., & Ven, P. van de (2004). Type of discharge and risk of recidivism measured by the HCR-20: A retrospective study in a Dutch sample of treated forensic psychiatric patients. *International Journal of Forensic Mental Health, 3*(2), 149-165. doi: 10.1080/14999013.2004.10471204

Walters, G. D. (2003). Predicting institutional adjustment and recidivism with the Psychopathy Checklist factor scores: A meta-analysis. *Law and Human Behavior, 27*(5), 541-558.

Walters, G. D., Duncan, S. A., & Mitchell-Perez, K. (2007). The latent structure of psychopathy: A taxometric investigation of the Psychopathy Checklist Revised in a heterogeneous sample of male prison inmates. *Assessment, 14*(3), 270-278. doi: 10.1177/1073191107299594

Walters, G. D., Knight, R. A., Grann, M., & Dahle, K. P. (2008). Incremental validity of the Psychopathy Checklist facet scores: Predicting release outcome in six samples. *Journal of Abnormal Psychology, 117*(2), 396-405. doi: 10.1037/0021-843X.117.2.396

Walters, G. D., Marcus, D. K., Edens, J. F., Knight, R. A., & Sanford, G. M. (2011). In search of the psychopathic sexuality taxon: Indicator size does matter. *Behavioral Sciences & the Law, 29*(1), 23-39. doi: 10.1002/bsl.964

Wayland, K., & O'Brien, S. (2014). Deconstructing antisocial personality disorder and psychopathy: A guideline-based approach to prejudicial psychiatric labels. *Hofstra Law Review, 42*, 519-588.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence* (Version 2). Burnaby, BC: Simon Fraser University.

Yang, M., Wong, S. C., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin, 136*(5), 740–767. doi: 10.1037/a0020473

Reconviction and revocation rates in Flanders after medium security treatment

International Journal of Law and Psychiatry, 47, 45-52

10

Jeandarme, I., Habets, P., Oei, T. I., & Bogaerts, S.

ABSTRACT

Aim. To examine the criminal outcome of Flemish forensic psychiatric patients (internees) after medium security treatment. Also, the effect of conditional release on recidivism of two subgroups (internees under conditional release and internees who received unconditional release) was examined.

Method. Reconviction rates and revocation rates were collected for all participants. Kaplan–Meier survival analyses were used to investigate recidivism rates while controlling for time at risk.

Results. During the 10-year period, 502 offenders were discharged from medium security treatment. Over a follow-up period averaging 3.6 years, 7.4% of discharged patients were reconvicted or received a new not guilty by reason of insanity (NGRI) verdict for a violent offense. One-quarter of the population had their conditional release revoked. Part of the study population was granted unconditional release. Reconviction rates were higher after unconditional release in comparison to conditional release.

Conclusions. The results of this study suggest that the court supervision of NGRI patients in Flanders is effective in protecting the community from further offending.

INTRODUCTION

Treatment outcome in forensic mental health is best measured over a broad range of areas, including clinical and humanitarian ones (Yiend et al., 2011). However, the prevention of future criminal behavior is the most important goal in forensic psychiatric treatment (Menghini, Ducro, & Pham, 2005). Different types of recidivism have been studied in offenders who were found not guilty by reason of insanity (NGRI), such as reconvictions, re-arrest, revocation and (re)incarceration rates and self-report (Heilbrun & Griffin, 1993). Reconviction rates underestimate the real size of recidivism but are considered to be a reliable measure of recidivism (Wartna, 2009). In the current study, general recidivism and violent recidivism were examined. General recidivism refers to reconvictions regarding any type of crime; violent recidivism refers to reconvictions associated with (sexual) violent reoffending.

It is difficult to determine whether recidivism rates are consistent with the success or failure of a forensic treatment because it is difficult to relate treatment results directly to recidivism as a number of factors during time at risk can influence individuals. In adult forensic populations, as far as we know, no meta-analyses show clear consistent associations between forensic treatment and a reduction in recidivism. In a research synthesis by Morgan et al. (2012) treatment effects of service providers to offenders with mental illness were examined across studies. Some studies suggested that forensic interventions can reduce symptoms of distress and improve offender's ability to cope with their problems, resulting in adapted behavioral markers such as institutional adjustment. Another meta-analysis mentioned a positive effect of interventions in terms of reducing continued criminal justice system involvement of any kind (e.g., decrease in time spent in detention and arrests after treatment). A similar positive trend was found for number of new convictions. In addition, violation of conditions appeared to be negatively correlated to treatment. Larger effect sizes were found for interventions comprising both an institutional and community component and some degree of voluntariness (Martin, Dorken, Wamboldt, & Wootten, 2012).

Regarding the Risk–Need–Responsivity principles (Andrews et al., 1990) and the Good Lives Model (Andrews, Bonta, & Wormith, 2011; Ward & Stewart, 2003), offenders can be divided into low, medium and high risk offenders depending on their treatment and criminogenic needs (low,

medium and high care), level of risk and protective factors (low, medium and high risk) and responsivity (degree of connection in the treatment) (Schuringa, Spreen, & Bogaerts, 2014). By weighting these three principles, judges can decide what the most suitable level of security is for offenders (low, medium and high security). Medium security units typically refer to a medium risk and security level according to its environmental, relational and procedural security characteristics. However, deciding whether an offender is eligible for one of the three levels of security remains arbitrary. Despite attempts, objective criteria to determine which setting is most appropriate for which type offender are lacking (e.g., Collins & Davies, 2005).

Medium security treatment and recidivism rates

International studies on (medium security) forensic psychiatric treatment have presented a mixed picture of recidivism rates that seems to vary from 7.1% to 63% for general recidivism and from 1.8% to 46% for violent recidivism over different follow-up periods of 1 to 10.8 years in NGRI populations (for an overview, see Hayes, Kemp, Large, & Nielssen, 2014) (studies not reported in Hayes et al., 2014; Edwards, Steed, & Murray, 2002; Friendship, McClintock, Rutter, & Maden, 1999; Harris, Rice, & Cormier, 2002; C. Lund, Hofvander, Forsman, Anckarsater, & Nilsson, 2013; Maden, Scott, Burnett, Lewis, & Skapinakis, 2004; Müller-Isberner, Freese, Jockel, & Gonzalez Cabeza, 2000; Nilsson, Wallinius, Gustavson, Anckarsäter, & Kerekes, 2011; Nowak & Nugter, 2014; Seifert & Moller-Mussavi, 2005; Tabita, de Santi, & Kjellin, 2012). To the best of our knowledge, no meta-analysis of recidivism rates in NGRI patients is currently available. In their meta-analysis of mainly American studies, Bonta, Blais, & Wilson (2014) found 39% general recidivism and 23% violent recidivism during a follow-up period of 4.9 years in a more heterogeneous group of offenders subjected to mental health intervention.

In Flanders, reconviction rates for NGRI acquittees are scarce and incomplete. In the southern part of Belgium (Wallonia), recidivism rates in medium and high risk offenders ranged from 21.4% to 34.4% for general recidivism and from 5.2% to 17.4% for violent recidivism after a follow-up period ranging from 2.5 to 4.2 years (Ducro & Pham, 2006; Menghini et al., 2005; T. H. Pham & Ducro, 2008; T. H Pham, Ducro, Marghem, & Réveillère, 2005). A nationwide Belgian study on reimprisonment after release from prison revealed a very high percentage (62.3%) of re-imprisonment in NGRI acquittees after 5.7 to 8.7 years (Robert & Maes, 2012).

Different insanity acquittee systems have been described in the literature (e.g., Dirks-Linhorst & Linhorst, 2006), most of them primarily focusing on public safety as their primary goal. Key components of the conditional release process of NGRI acquittees include the development and monitoring of conditions of release and access to revocation and inpatient hospitalization when violations of conditions occur (Dirks-Linhorst & Linhorst, 2006). In the context of risk management and the prevention of recidivism, most conditionally released individuals are required to follow treatment and (probation) supervision. Not adhering to prescribed rules and ancillary conditions often results in a return to a secure, inpatient facility for further treatment and/or confinement. Therefore, in NGRI populations, typically two outcome metrics related to *failure* are being used: the acquisition of new criminal charges and/or conditional release revocation due to criminal acts or rule violations. Literature demonstrates that revocations for rule violations are higher than revocations for acquisition of new criminal charges (Vitacco, Vauter, Erickson, & Ragatz, 2014; Wiederanders, 1992). Revocation rates of rule violations range from 5% to 49% (Bertman-Pate et al., 2004; Green et al., 2014; Manguno-Mire, Coffman, DeLand, Thompson, & Myers, 2014; Vitacco et al., 2008; Wiederanders, Bromley, & Choate, 1997) over different follow-up periods ranging from 1.7 to 5.1 years.

Conditional release and monitored aftercare programs following intramural treatment are considered to be effective to ensure safe transitions from secure facilities to community life. However, regimens of treatment and supervision are seldom reported or quantified, and studies of the effectiveness of conditional release programs are difficult to compare (Wiederanders et al., 1997). Although there has been past research confirming that post-release supervision and community treatment can reduce recidivism (e.g., Wiederanders, 1992), a systematic review shows that little empirical evidence exists to conclude that long term supervision remains effective (van Gestel, van der Knaap, & Hendriks, 2006). In addition, studies have shown that recidivism can be reduced by implementing (forensic) ambulatory care after release (i.e., Home Office restrictions requiring patients to accept supervision and treatment following discharge: Coid, Hickey, Kahtan, Zhang, & Yang, 2007; specialised forensic outpatient clinics: Schmidt-Quernheim & Seifert, 2013).

Current study

This study was conducted among offenders found not criminally responsible for a committed crime (in Belgium referred to as internees) and focused on recidivism after treatment in a medium security unit (MSU). First, a brief background about Belgian legislation and practice is given because it differs from most countries.

Legislative background

Under Belgian law, internment is a safety measure imposed by a (investigating) judge to an offender if the latter is found not guilty by reason of insanity (NGRI). Offenders can be interned if it is proven that they have committed an offense[1] and they are found irresponsible or severely diminished responsible at the moment of the trial as a consequence of either a status of insanity or a serious mental deficiency which makes the person unable to (fully) control his actions. While in most cases a psychiatrist (and psychologist) will perform a forensic psychiatric evaluation to determine criminal responsibility, this is not mandatory, nor does a common law standard for legal insanity exists. Internment is not a punishment, nor can it be combined with a criminal sanction. It is an indefinite safety measure aiming to prevent (further) harm to society and provide treatment for the internee (Goethals, 1997). On a Belgian population of about eleven million inhabitants, about 300 to 400 people are annually placed under this internment measure (Department of Criminal Justice Policies, 2012). Over the years, the number of internees has been rising; at the end of 2013, there were about 3,820 internees in Belgium (Deckers et al., 2014).

A multidisciplinary court chaired by a judge, the Commission of the Protection of Society (CPS), is responsible for the implementation of the internment. While the prosecutor advices the court, only the other members of the CPS (psychiatrist, lawyer and judge) decide in which type of setting the internee will be treated and when he or she will be conditionally or unconditionally released. Automatic hospitalization is not required at the time of acquittal since conditional release into the community is also an option. According to the specific treatment needs (low, medium or high care), risk of recidivism (low, medium or high risk) and security level (low, medium or high security) assessed by the mental health probation officer or the psychosocial prison team, internees in theory can either reside in prison, or in forensic psychiatric units, regular psychiatric units or even protected houses or the community receiving ambulatory care.

However, forensic beds were not implemented in Flanders until 2001, when the first medium security units emerged and only recently a high security forensic hospital opened (FPC Ghent since the end of 2014). As a consequence, many Flemish internees (1087 in 2013) deemed too dangerous for community supervision still remain in prison without adequate treatment (Deckers et al., 2014; Moens & Pauwelyn, 2012; Vandevelde et al., 2011). Every six months, the internee can appear before the CPS to ask for his or her conditional release. When an internee is treated outside the prison system, the internee is conditionally released under the authority of the CPS. On conditional release, the patient's liberty is dependent on their adhering to several requirements, usually including living in a particular treatment setting and continuing to receive psychiatric supervision and treatment. As soon as the security level has diminished, the internee will either be referred conditionally to a low security (inpatient) facility or released into the community with outpatient care. All treatment options are coupled with a judicial mandate to receive treatment and follow all specified rules. Because of breach of conditions, the internee can be readmitted to prison, and his conditional release will be revoked. As soon as the internee is no longer a risk for society, the CPS will release the internee unconditionally and stop the internment measure. After that, the CPS will have no authority over him/her anymore. A forensic psychiatric examination prior to unconditional release is not mandatory.

The current study focused on general and violent recidivism rates in (conditionally) released internees discharged from medium security forensic psychiatric settings. In this study, which is the first in its kind in Flanders, reconviction data and re-imprisonment data were examined. In order to examine the effect of conditional release on recidivism, reconviction data of two subgroups (internees under conditional release and internees who received unconditional release) were compared. Furthermore, in the unconditional release group, a comparison between the conditional release period compared to the unconditional release period was made. The following hypotheses were investigated:

1. Post-treatment reconviction rates[2] will be lower than pre-treatment reconviction rates.
2. Reconviction rates will be lower during conditional release when compared to unconditional release (i.e., CPS supervision will have a positive effect on reconviction rates).

3. Post-treatment reimprisonment rates due to conditional release revocation will be higher than reconviction rates.

MATERIAL AND METHODS

Setting and study population

The study was conducted at the three forensic medium security units in Flanders, Belgium. These units focus on psychotic and personality disordered internees. Rarely, an internee with a primarily substance use disorder, intellectual disability or paraphilia is accepted for treatment, although all of these conditions can be present as comorbid disorders (see Boers, Vandevelde, Soyez, De Smet, & Ting To, 2011 for a description of treatment programs). In the period 2001–2010, 542 internees under the authority of a Flemish CPS were referred to one of the forensic units. Eleven participants refused to participate in the study. Five participants died during their treatment, and 24 were not discharged from the medium security unit at the end of the study. The study population therefore consisted of 502 internees supervised by CPS Antwerp (45.2%, *n* = 227), CPS Ghent (38.2%, *n* = 192), CPS Leuven (13.1%, *n* = 66) or CPS Vorst (3%, *n* = 15)[3] and discharged during the study period after forensic treatment in three Flemish cities with medium security units: Bierbeek (28.1%, *n* = 141), Rekem (30.1%, *n* = 151) or Zelzate (41.8%, *n* = 210). All internees were examined from date of discharge from the medium security hospital to the end of the study period (December 31, 2010) or date of death. The observation period included different degrees of supervision while conditionally released, as the internee could be transferred between institutional and ambulatory care, or back to the institution or prison, either for relapse in offenses, breach of conditions or worsening of the psychiatric condition. After unconditional release from the internment measure, judicial conditions are no longer imposed. A minority of the population (13.7%, *n* = 69) was initially conditionally released and later on unconditionally released form the internment measure during the study period (Figure 1).

The majority of the discharged population (93.2%, *n* = 468) was referred to a medium security unit directly from the prison system, while a minority was referred from low security settings. Mean age at first medium security discharge was 37.8 years (*SD* = 10.9, range = 18.8–73.5). The majority was male (95.2%, *n* = 478) and born in Belgium[4] (89.6%, *n* = 447). The mean

duration of stay in the medium security unit was 675.9 days (*SD* = 510.1 days, range = 8–2729 days).

Most participants (81.5%, *n* = 402)[5] had previously been admitted to general psychiatric wards and were previously convicted and/or interned (84.1%, *n* = 422). Most common (principal) index offenses were violent offenses (77.5%, *n* = 389) and property offenses (18.5%, *n* = 93). Other index offenses were drug-related offenses (2.2%, *n* = 11), other offenses (1%, *n* = 5) and non-violent sexual offenses (0.8%, *n* = 4). The most common DSMIV-TR diagnoses were personality disorder (70.7%, *n* = 355), substance use disorder (56.2%, *n* = 282) and psychotic disorder (43.8%, *n* = 220). Overlap between DSM diagnoses was frequent; 71.3% of the internees (*n* = 358) had combined Axis I and Axis II pathology.

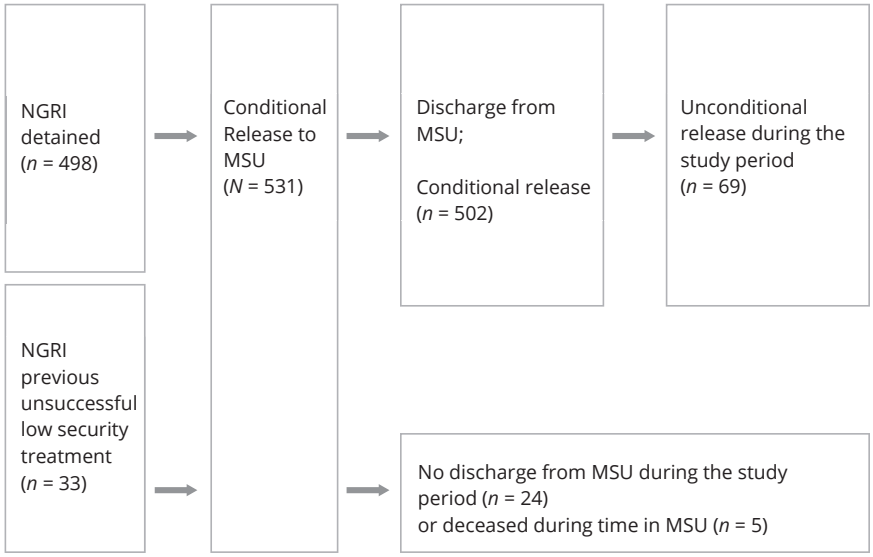


Figure 1. Course of internment measure.

Baseline and follow-up measures

The following baseline variables were obtained from multiple sources (medium security unit files and/or CPS files):

- Age at discharge: age at (first) discharge from MSU.

- Violent index offense: the primary index offense involves at least one (sexual) violent offense.
- Previous criminality: age at the first registered sentence[6] as well as the total number of previous sentences (including the index internment measure).
- Diagnoses: all (comorbid) diagnoses were made on the basis of extensive file information in accordance with the Diagnostic and Statistical Manual of mental disorders-IV-Text Revision (DSM-IV-TR; American Psychiatric Association, 2000) by the first author after consulting with the treating psychiatrists.
- Intelligence: scores on the Wechsler Adult Intelligence Scale-III (WAIS-III; Wechsler 2005), found in the files, were used.
- Psychopathy: scores on the Psychopathy Checklist-Revised (PCL-R; Vertommen, Verheul, de Ruiter, & Hildebrand, 2002), found in the files, were used.

Recidivism data based on official sentences were described for the whole group and for the conditionally released and unconditionally released groups separately. All the recidivism data were based on criminal records, collected from the Central Criminal Records of the Ministry of Justice. In addition, dates and specifications of offenses were retrieved from the different court administrations. Furthermore, recidivism data based on revocations – and hence detentions – were described. Detention periods were extracted from the prison registration system SIDIS (*Système informatique de Détention/Detentie Informatie Systeem*). The follow-up time was classified into two categories – conditional and unconditional release – depending on the CPS supervision. The following follow-up variables were obtained from abovementioned multiple sources:

- General recidivism: number of recidivists in all types of crimes (violent and non-violent, including traffic offenses) that resulted in a new sentence (either reconvictions or new internment measures) and the total number of new sentences.
- Violent recidivism: number of recidivists in violent crimes that resulted in a sentence and the total number of violent crimes, resulting in a new sentence. Violent crimes were defined as any physical assault or serious threat to the victim including sexually violent offenses (e.g., assault, robbery with assault, threats and sexual hands-on offenses).

- Detention periods: number of participants that served time in prison, as well as the number of detentions and the total duration of prison sentences during the study period.
- Date of unconditional release: date of termination of the internment measure.
- Conditional release period: days from date of first discharge from medium security unit till unconditional release from the internment measure or end of data collection (or time of death) minus number of days spend in re-admission. Conditional release periods include, e.g., ambulatory care while living in the community or residing in a facility with lower security such as a general psychiatric ward and detention periods.
- Unconditional release period: days from date of unconditional release till end of study (or time of death).

Ethics

The study was approved by the Ethics committee of Antwerp University Hospital on January 24, 2011.

Statistical analyses

Analyses were conducted using the software package SPSS version 22 (IBM Corp, Released 2013). Chi^2 or Fisher's exact analyses were used to investigate differences between two categorical variables and McNemar analyses in case of paired samples. The Mann–Whitney U test or the Kruskall–Wallis test was used for continuous variables, and Wilcoxon signed ranks test was used in case of paired samples. The Kaplan–Meier estimator was used for the survival analyses. The percentage of released patients from their date of conditional release until the particular event studied was calculated. The events studied included date of first general re-offense and first violent re-offense. Survival analyses were repeated for patients who were conditionally released and patients who were unconditionally released. Characteristics and recidivism rates of the conditionally released group (n = 433) were compared to the unconditionally released group (n = 69). Finally, a temporal comparison was performed: in the group of patients who had been unconditionally released (n = 69), time at risk and reconviction rates were compared during the

time they were conditionally released with the time they were unconditionally released. For each participant, a comparison was made between the annualized rate of sentences pre- and post-treatment to assess the evolution in criminal behavior according to the following formula: (Gilleir, Easton, & Ponsaers, 2010):

$$\text{Density sentences pre-treatment} = \frac{\text{number of sentences before admission}}{\text{Time period (18 years until first admission)}} \times 365$$

$$\text{Density sentences post-treatment} = \frac{\text{number of sentences after admission}}{\text{Time period (first discharge until 31/12/2010 or death)}} \times 365$$

RESULTS

Reconviction rates

Sixty-three (12.5%) of the 502 internees were reconvicted for an offense at least once. In the group of the recidivists, 107 reconvictions for 185 offenses were examined: violent offenses (38.4%, *n* = 71), property offenses (13.5%, *n* = 25), drug offenses (3.8%, *n* = 7), sexual non-violent offenses (1%, *n* = 2), and other offenses (43.2%, *n* = 80), consisting of mainly traffic offenses (*n* = 61). Thirty-seven internees (7.4%) had reoffended in a violent offense. Violent offenses consisted of attempted homicide (*n* = 1), arson (while people were present in building) (*n* = 3), sexual assault (*n* = 3), burglary with assault or battery (*n* = 13), threats (*n* = 13) and other violent offenses (*n* = 38). The mean duration to the first reconviction was 1009.8 days (*SD* = 782.1, range = 37–2920) for the whole group (see Figure 2 for survival analyses).

The mean time at risk period was 1332.5 days (*SD* = 807.4, range = 7–3426 days). There was no difference in general (*p* = .79) or in violent recidivism between the different CPS (*p* = .79).

Overall, when comparing sentence density pre- and post-treatment, a significant decrease in sentence density was found (*p* = .00). In 5.2% of the population (*n* = 26), an increase in sentence density was found after medium security treatment, whereas a decrease was observed in the majority of the population (94.8%, *n* = 476).

While all internees (*N* = 502) were initially conditionally released for a period of time (*M* = 1193.1 days, *SD* = 740.6, range = 7–3426), only a minority (*n* = 69) had later on a follow-up time after unconditional release (*M* = 1014.4 days, *SD* = 564.5, range = 44–2279). In the group that was

eventually unconditionally released (*n* = 69), recidivists were more common after MSU treatment than in the group remaining under conditional release in the whole study period (*n* = 433).

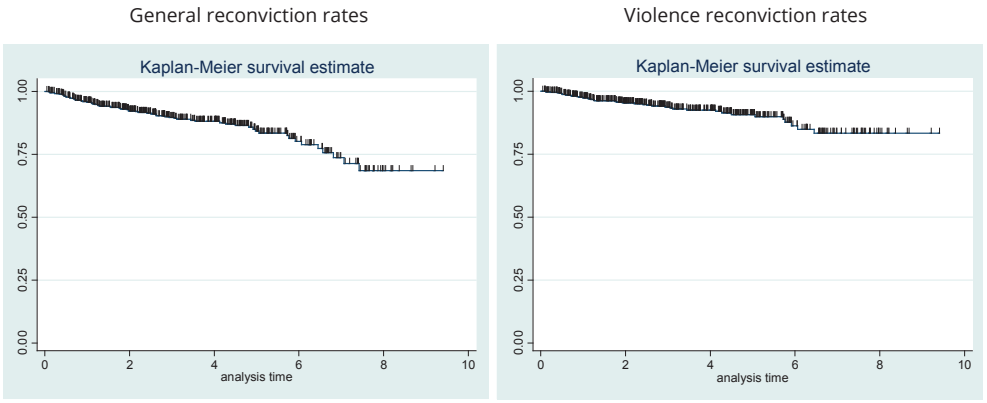


Figure 2. Kaplan-Meier survival curve displaying the proportion of patients surviving (y-axis) over time (years) (x-axis). Censoring is indicated by the vertical lines.

Reconviction rates compared between conditional and unconditional release groups

In the former, the proportion of recidivists (27.5%, *n* = 19) was more than double than in the conditional group (10.2%, *n* = 44) concerning all types of crimes (*p* = .00). For violent types of crimes, the increase was not significant (13%, *n* = 9 vs. 6.5%, *n* = 28) (*p* = .05).

Survival analyses showed no differences between reconviction rates comparing conditional release and unconditional release subjects (see Figure 3) (*p* = .12, violence *p* = .76).

In Table 1, recidivism rates during the conditional time at risk period are compared with recidivism rates during the unconditional time at risk period within the internees that have been unconditionally released (*n* = 69). The increase in general (20.3% vs. 10.1%) and in violent recidivism (10.1% vs. 2.9%) was not significant. In the group of the 14 recidivists after unconditional release, 20 reconvictions for 33 offenses were examined: violent offenses (42.4%, *n* = 14), property offenses (3.0%, *n* = 1) and other offenses (54.5%, *n* = 18), consisting of mainly traffic offenses (*n* = 13). Seven internees reoffended in a violent offense. Violent offenses consisted of burglary with assault or battery (*n* = 1), threats (*n* = 2) and other violent offenses (*n* = 11).

Comparisons of criminal history, index offenses and clinical features between internees who were still under CPS conditions at the end of the study and internees that were unconditionally released are presented in Table 2. Some baseline differences were noted. Unconditionally released internees were older at first conviction (p = .01), had less previous convictions (p = .01), less absconding during MSU treatment (p = .00), lower PCL-R scores (p = .01) and less often a personality disorder (p = .03) or an antisocial personality disorder (p = .01).

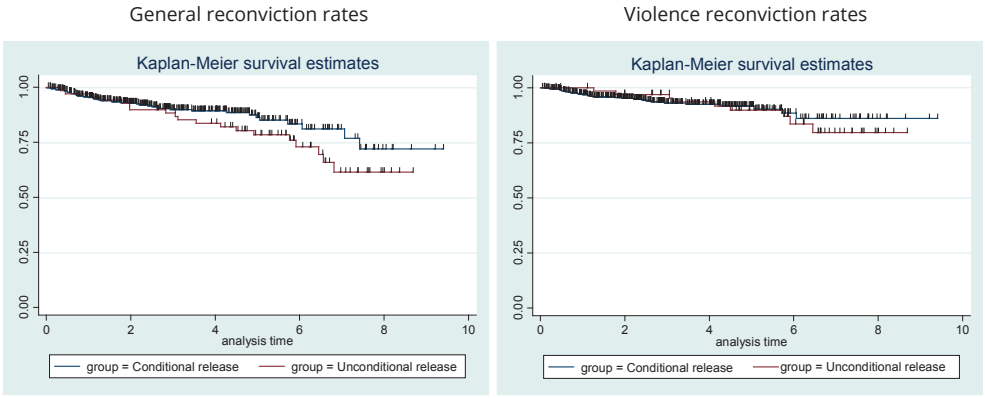


Figure 3. Kaplan-Meier survival curves comparing survival rates between patients who received an unconditional release with patients who did not receive unconditional release. Proportion of surviving is displayed on the y-axis. Time since discharge from medium security (years) is displayed on the x-axis. Censoring is indicated by the vertical lines.

Revocation rates

Conditional release was revoked in a quarter of the population (26%, n = 130). As a consequence, these 130 patients were re-incarcerated in prison while a minority was re-incarcerated in the unconditional release period (2.9%, n = 2). The average number of years of reimprisonment during conditional release was 1.5 years (SD = 1.6 years, range = 1 day–8.5 years), and the average number of reimprisonments was 1.4 (SD = 0.9, range = 1–6). There was no difference in

Table 1. Comparison of reconviction rates of patients with unconditional release during conditional and unconditional release

	During conditional release	During unconditional release	*p*
Time at risk			
Mean number of days (*SD*)	1139.8 (582.5) Range = 334-2472	1014.4 (564.5) Range = 44-2279	.62
Recidivists (PB) n = 69			
All offenses *n* (%)	7 (10.1)	14 (20.3)	.14
Violent offenses *n* (%)	2 (2.9)	7 (10.1)	.18
Convictions (EB)			
All offenses	15	20	.39
Violent offenses	2	9	.08

Note. PB = person based; EB = event based.

Table 2. Comparison between patients with and without unconditional release

	All offenders (*n* = 502)	No unconditional release (*n* = 433)	Unconditional release (*n* = 69)	*p*
Criminal history (including index)				
Age at first conviction *M* (*SD*)	24.7 (9.1)	24.2 (8.8)	27.8 (10.7)	.01**
Previous convictions				
Violent offense *M* (*SD*)	2.4 (1.9)	2.4 (2.0)	2 (1.6)	.17
All offenses *M* (*SD*)	6.2 (5.8)	6.5 (5.9)	4.7 (4.4)	.01*
Juvenile history yes/no *n* (%)[1]	195 (41%)	176 (43.2%)	19 (27.5%)	.01*
Violent index offenses *n* (%)	389 (77.5%)	332 (76.7%)	57 (82.6%)	.27
Medium security treatment				
Absconding history *n* (%)	153 (30.5%)	143 (33%)	10 (14.5%)	.00**
Age at first discharge *M* (*SD*)	37.8 (10.9)	37.5 (10.7)	39.9 (11.9)	.13
Clinical history				
WAIS-III scores *M* (*SD*)[2]	80.5 (17.1)	80.9 (17.4)	77.5 (14.6)	.39
PCL-R scores *M* (*SD*)[3]	21.2 (6.8)	21.6 (6.5)	16.8 (8.2)	.01*
Psychiatric diagnoses n (%)				
Psychosis	220 (43.8%)	191 (44.1%)	29 (42%)	.75
Any personality disorder	355 (70.7%)	314 (72.5%)	41 (59.4%)	.03*
Antisocial personality disorder	127 (25.3%)	119 (27.5%)	8 (11.6%)	.01*
Substance use disorder	282 (56.2%)	247 (57%)	35 (50.7%)	.33

[1] 5.2% missing; [2] 46.1% missing; [3] 64.1% missing.
* $p < 0.05$. ** $p < 0.001$.

revocation rates between the different CPS (p = .96), or a difference in mean duration of re-imprisonment (p = .61). In addition, re-imprisonment rates (26%, n = 130) during the conditional period were higher than reconviction rates (10.2%, n = 51)[7] (p = .00).

DISCUSSION

This paper contributes to the literature in several ways. First, recidivism rates in a Flemish forensic population – internees – were examined for the first time. As such, the study adds to the literature about the outcome of NGRI offenders, which is still scarce (Hayes et al., 2014). Second, describing how the judicial system in Flanders handles NGRI acquittees can be helpful to professionals in other countries. It was examined whether and how CPS supervision had a positive effect on reconviction rates by comparing recidivism rates during conditional and unconditional release. Lastly, post-treatment re-imprisonment rates were compared to the reconviction rates.

Recidivism rates were reassuringly low; 12.5% of the internees recidivated with a general offense (including fines for traffic offenses) and 7.4% with a violent offense. In line with a recent retrospective study by (Hayes et al., 2014), none of the unconditionally released patients reoffended in a serious crime. Furthermore, a decrease was measured between pre- and post-treatment offending and thus confirming the first hypothesis. Although these results are encouraging, without a control group, it remains unclear whether they can be attributed to a treatment effect. Demographic features could not explain the relativity low recidivism rates. For example, as women tend to have lower recidivism rates than men (Maden, Skapinakis, Lewis, Scott, & Jamieson, 2006), higher recidivism rates in a study using a predominantly male population (95%) were expected. Furthermore, other criminogenic features such as ages at release, amount of patients with a personality disorder (71%) as well as the prior offense history, were at least comparable with most other studies. Comparisons between international studies are however difficult due to differences in several core aspects, such as the definition of the population, follow-up periods, outcome measures, settings and legal regulations (Lund, Forsman, Anckarsäter, & Nilsson, 2012). With regard to differences in outcome measures, most international studies investigating recidivism used reconviction as outcome measure, while some

studies used re-arrest (Bertman-Pate et al., 2004) or an even broader definition of recidivism such as re-admission in a forensic hospital or any police contact (Lee, 2003). The more stringent definitions will logically result in lower recidivism rates. The definition of violent offenses in studies is not always clearly described. In the current study a broad definition of interpersonal violence was used while some international studies used more strict definitions excluding verbal violence, property offenses with violence or sexual violence (e.g., Coid et al., 2007). Furthermore, the setting and type of supervision can differ after forensic treatment and change over time. It can be expected, for example, that during hospitalization recidivism rates will be lower. In studies, there is wide variation in the reported rates of rehospitalization, and it is likely that there is some confusion between hospital re-admission and conditional release revocation. Some studies include periods of re-admission and/or incarceration (e.g., Friendship et al., 1999; Lund et al., 2013), while other studies exclude all periods of residential stay (Coid et al., 2007). In this study, rehospitalization in MSU was removed from the analyses, but other – general psychiatric – hospitalizations could not be traced.

Support was partly found for the second hypothesis. When comparing the unconditional release group (*n* = 69) to the conditional group (*n* = 433), there were significantly more recidivists with a general offense in the unconditional release group than in the conditional release group. The difference was trend significant for violent recidivism. When investigating the moment of recidivism in the unconditional release group, it became apparent that violent and general recidivism rates were higher after unconditional release in comparison with conditional release for general, and for violent recidivism. However, this difference did not reach statistical significance. In addition, it was investigated on which factors the internees with an unconditional release differed from the internees still under conditional release. Internees with an unconditional release were older at first offense, had committed fewer offenses, had lower PCL-R scores and were less likely to be diagnosed for (antisocial) personality disorder or have absconded during treatment. How evaluators decide to release NGRI patients conditionally or unconditionally is largely unknown. The study by Gowensmith, Bryant, and Vitacco (2014) investigated the procedures used by 89 forensic evaluators from nine states of the United States in conditional release readiness evaluations. The study showed that not only was there a lack of uniformity on any aspect of the decision making processes, evaluators de-emphasized important

factors that have been empirically proven to be related to recidivism. Given the results from the present study, it can be stated that the CPS granted unconditional release to internees in which the risk of reoffending could be regarded as low based on abovementioned criminogenic characteristics. Contradictory, it was this group who eventually had higher recidivism rates than the conditionally released group. This could be explained by the fact that after unconditional release, supervision is no longer mandatory, as has been shown in other research. For example, a German study by Butz, Mokros, and Osterheider (2013) found significantly lower recidivism rates (2%) in high risk forensic psychiatric patients when these patients received specialized aftercare when compared to patients without aftercare (i.e., unconditionally released) (18.5%). In this study, post-treatment revocation rates were significantly higher than reconviction rates and confirming the third hypothesis. In Belgium, revocation results in re-imprisonment. Re-imprisonment rates in this study were lower than found in a national study on reimprisonment, but with longer follow-up periods and comprising internees with and without forensic treatment (Robert & Maes, 2012).

Furthermore, our findings are in line with international research that found that revocation rates for violating conditions are higher as opposed to new criminal activity (Vitacco et al., 2008). Revocations, taking place in public protection-oriented conditional release programs, have been identified as a major cause of low re-offense rates. However, the relationship between revocations and re-offenses was not linear (Wiederanders et al., 1997). Other studies have suggested that conservative court decisions (e.g., cautious decisions regarding release, preventative hospitalizations) result in lower (violent) reconviction rates (Hayes et al., 2014; Luettgen, Chrapko, & Reddon, 1998; Skipworth, Brinded, Chaplow, & Frampton, 2006).

Clinical implications

The successful (conditional) release by CPS in Flanders has several key success features such as a centralized responsibility, uniform system of supervision by mental health probation officers, flexible procedure and a close collaboration with the forensic treatment facilities. Furthermore, there is some voluntariness in choosing a treatment option in collaboration with the internee. As was shown in the current study, this collaboration resulted in safe release plans. Policymakers

and clinicians should also be aware of the importance of (any form of) supervision. Namely, recidivism rates were higher when the CPS no longer had jurisdiction over the internee.

On the other hand, as is the case with the therapeutic approach, there remains an inherent tension between therapeutic potential and due process principles (Perlin, 2014). As was evident in the current study, revocation rates were high, and under the current law, re-incarceration occurs without due process. The situation of internees in prison has often been criticized and yielded the Belgian government several convictions. Currently, a new internment law is under revision. Further research is needed to evaluate the (criminal) outcome of this more legal rights-based model with more bureaucratic procedures and due process.

Treatment providers should be equally concerned with the high dropout rates. As has been shown in other research, treatment dropout can have dramatic effects in forensic populations (McMurran & Theodosi, 2007; Pham et al., 2005). In that respect, more attention should be paid to responsivity issues and treatment alliances in order to keep NGRI offenders in treatment (Gannon & Ward, 2014). In an attempt to facilitate admissions from penitentiary to psychiatric hospitals and to minimize revocations, the InReach pilot project was initiated (Stassen, Habets, Mertens, De Laender, & Jeandarme, 2014). The main idea of the project was to have a psychiatric nurse on the ward doing pretherapeutic and motivational activities on a regular basis in the penitentiary, forming a bridge between prison and hospital. A motivational approach is used to support these patients in making the transition from penitentiary to hospital. Namely, often problems arise early in the treatment process (potentially resulting in re-incarceration), which are likely to be avoided by proper preparation of patients prior to treatment (e.g., unrealistic expectations).

Limitations

Although the overall sample size was large, the recidivism rates were rather low, resulting in a reduction of statistical power. Type II errors can therefore not be ruled out in the analyses comparing recidivists with non-recidivists. This could explain the absence of significant differences in recidivism between, e.g., the conditional and unconditional group. Also, because of low recidivism rates, survival analyses investigating the effect of several risk factors on survival

(e.g., Cox regression analyses) could not be performed. Namely, Cox regression power analysis showed that 66 events were required to detect a 0.5 reduction in the hazard for a binary covariate of interest with standard deviation 0.5, using a two-sided 5% Wald test with a power of 80%. Furthermore, because of the skewness of the data, non-parametric tests have been used.

When using official files to determine recidivism rates, the problem of *dark numbers* needs to be taken into account. Dark numbers refer to the offenses that are not officially registered. Namely, a number of offenses are not reported to the police, and the reported offenses do not always result in charges, let alone in convictions. As a result, reconviction rates underestimate reoffending. However, the difference between reoffending and reconviction is less marked with more severe offending (Skipworth et al., 2006). Another issue to be aware of is that the follow-up period included episodes of incarceration and reconviction during detention, while periods of medium security hospitalization during the follow-up period were excluded.

CONCLUSIONS

The relatively low recidivism rate found in the current study could not be attributed to study population features (e.g., predominantly male, high proportion of personality disorder), suggesting that the interment procedure is successful in reducing the level of dangerousness for the community. This reduction of dangerousness is mostly apparent as long as the interment measure is in effect and the patient is under some form of supervision. Higher revocation rates were noted, emphasizing that balancing public protection and least restrictive alternatives for individual freedom remain at this point an important issue.

FOOTNOTES

[1] All offenses for which the Criminal Law sets a minimum penalty of at least eight days are included.

[2] Reconviction rates refer to new sentences, either new reconvictions or new internment measures.

[3] In addition in two cases the CPS file could not be traced.

[4] Three missings.

[5] Nine missings.

[6] Sentences included convictions as well as internment measures.

[7] Fines excluded the reconviction rate was 7.4% (*n* = 37).

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17*, 19-52.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The Risk-Need-Responsivity (RNR) model: Does adding the Good Lives Model contribute to effective crime prevention? *Criminal Justice and Behavior, 38*(7), 735-755. doi: 10.1177/0093854811406356

Bertman-Pate, L. J., Burnett, D. M. R., Thompson, J. W., Calhoun, C. J., Deland, S., & Fryou, M. (2004). The New Orleans forensic aftercare clinic: A seven year review of hospital discharged and jail diverted clients. *Behavioral Sciences & the Law, 22*(1), 159-169. doi: 10.1002/bsl.575

Boers, A., Vandevelde, S., Soyez, V., Smet, S. De, & To, W.T. (2011). Het zorgaanbod voor geïnterneerden in België. *Panopticon, 32*(2), 17-38.

Bonta, J., Blais, J., & Wilson, H. A. (2014). A theoretically informed meta-analysis of the risk for general and violent recidivism for mentally disordered offenders. *Aggression and Violent Behavior, 19*(3), 278-287. doi: 10.1016/j.avb.2014.04.014

Butz, M., Mokros, A., & Osterheider, M. (2013). Specific outpatient treatment for forensic patients after release. *Psychiatrische Praxis, 40*(1), 36-42. doi: 10.1055/s-0032-1327197

Coid, J., Hickey, N., Kahtan, N., Zhang, T., & Yang, M. (2007). Patients discharged from medium secure forensic psychiatry services: Reconvictions and risk factors. *The British Journal of Psychiatry, 190*, 223-229. doi: 10.1192/bjp.bp.105.018788

Collins, M., & Davies, S. (2005). The Security Needs Assessment Profile: A multidimensional approach to measuring security needs. *International Journal of Forensic Mental Health, 4*(1), 39-52.

Deckers, A., Seynnaeve, K., Smedt, A. De, Dheedene, J., Vandenplas, E., & Auwera, A. Van der (2014). *Geïnterneerden in detentie op 24/12/2013*. Unpublished report.

Department of Criminal Justice Policies. (2012). *Statistiek interneringen*. Retrieved from http://www.dsb-spc.be/

Dirks-Linhorst, P. A., & Linhorst, D. (2006). A description of the design and costs of an insanity acquittee conditional release, monitoring, and revocation system. *International Journal of Forensic Mental Health, 5*(1), 55-65. doi: 10.1080/14999013.2006.10471230

Ducro, C., & Pham, T. (2006). Evaluation of the SORAG and the Static-99 on Belgian sex offenders committed to a forensic facility. *Sexual Abuse: A Journal of Research and Treatment, 18*(1), 15-26. doi: 10.1007/s11194-006-9003-6

Edwards, J., Steed, P., & Murray, K. (2002). Clinical and forensic outcome 2 years and 5 years after admission to a medium security unit. *The Journal of Forensic Psychiatry, 13*(1), 68-87.

Friendship, C., McClintock, T., Rutter, S., & Maden, A. (1999). Re-offending: Patients discharged from a regional secure unit. *Criminal Behaviour and Mental Health, 9*(3), 226-236. doi: 10.1002/cbm.315

Gannon, T. A., & Ward, T. (2014). Where has all the psychology gone? A critical review of evidence-based psychological practice in correctional settings. *Aggression and Violent Behavior, 19*, 435-446. doi: 10.1016/j.avb.2014.06.006

Gestel, B. van, Knaap, L. M. van der, & Hendriks, A. (2006). *Toezicht buiten de muren: Een systematische review van extramuraal toezicht op TBS-gestelden en vergelijkbare groepen in het buitenland*. Den Haag: Wetenschappelijk Onderzoek- en Documentatiecentrum Justitie.

Gilleir, F., Easton, M., & Ponsaers, P. (2010). Een U-Turn voor jongvolwassen veelplegers. *Panopticon, 31*(1), 5-19.

Goethals, J. (1997). De wet tot bescherming van de maatschappij in een historisch perspectief. In J. Casselman, P. Cosyns, J. Goethals, M. Vandenbroucke, D. De Doncker, & C. Dillen (Eds.), *Internering* (pp. 11-38). Leuven/Apeldoorn: Uitgeverij Garant.

Gowensmith, W. N., Bryant, A. E., & Vitacco, M. J. (2014). Decision-making in post-acquittal hospital release: How do forensic evaluators make their decisions? *Behavioral Sciences & the Law, 32*(5), 596-607. doi: 10.1002/bsl.2135

Green, D., Belfi, B., Griswold, H., Schreiber, J. M., Prentky, R., & Kunz, M. (2014). Factors associated with recommitment of NGRI acquittees to a forensic hospital. *Behavioral Sciences & the Law, 32*(5), 608-626. doi: 10.1002/bsl.2132

Harris, G. T., Rice, M. E., & Cormier, C. A. (2002). Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior, 26*(4), 377-394.

Hayes, H., Kemp, R. I., Large, M. M., & Nielssen, O. B. (2014). A 21-year retrospective outcome study of New South Wales forensic patients granted conditional and unconditional release. *The Australian and New Zealand Journal of Psychiatry, 48*(3), 259-282. doi: 10.1177/0004867413507610

Heilbrun, L., & Griffin, P. A. (1993). Community-based forensic treatment of insanity acquittees. *International Journaf of Law and Psychiatry, 16*, 133-150.

IBM Corp. (Released 2013). *IBM SPSS Statistics for Windows, Version 22.0*. Armonk, NY: IBM Corp.

Lee, D. T. (2003). Community-treated and discharged forensic patients: An 11-year follow-up. *International Journal of Law and Psychiatry, 26*(3), 289-300. doi: 10.1016/s0160-2527(03)00039-6

Luettgen, J., Chrapko, W. E., & Reddon, J. E. (1998). Preventing violent re-offending in Not Criminally Responsible patients: An evaluation of a continuity of treatment program. *International Journal of Law and Psychiatry, 21*(1), 89-98. doi: 10.1016/S0160-2527(97)00051-4

Lund, C., Forsman, A., Anckarsäter, H., & Nilsson, T. (2012). Early criminal recidivism among mentally disordered offenders. *International Journal of Offender Therapy and Comparative Criminology, 56*(5), 749-768. doi: 10.1177/0306624X11411677

Lund, C., Hofvander, B., Forsman, A., Anckarsater, H., & Nilsson, T. (2013). Violent criminal recidivism in mentally disordered offenders: A follow-up study of 13-20 years through different sanctions. *International Journal of Law and Psychiatry, 36*(3-4), 250-257. doi: 10.1016/j.ijlp.2013.04.015

Maden, A., Scott, F., Burnett, R., Lewis, G. H., & Skapinakis, P. (2004). Offending in psychiatric patients after discharge from medium secure units: Prospective national cohort study. *British Medical Journal, 328*(7455), 1534. doi: 10.1136/bmj.38075.467569.EE

Maden, A., Skapinakis, P., Lewis, G., Scott, F., & Jamieson, E. (2006). Gender differences in reoffending after discharge from medium-secure units. National cohort study in England and Wales. *The British Journal of Psychiatry, 189*, 168-172. doi: 10.1192/bjp.bp.105.014613

Manguno-Mire, G. M., Coffman, K. L., DeLand, S. M., Thompson, J. W., Jr., & Myers, L. (2014). What factors are related to success on conditional release/discharge? Findings from theNew Orleans Forensic Aftercare Clinic: 2002-2013. *Behavioral Sciences & the Law, 32*(5), 641-658. doi: 10.1002/bsl.2138

Martin, M. S., Dorken, S. K., Wamboldt, A. D., & Wootten, S. E. (2012). Stopping the revolving door: A meta-analysis on the effectiveness of interventions for criminally involved individuals with major mental disorders. *Law and Human Behavior, 36*(1), 1-12. doi: 10.1037/h0093963

McMurran, M., & Theodosi, E. (2007). Is treatment non-completion associated with increased reconviction over no treatment? *Psychology, Crime & Law, 13*(4), 333-343. doi: 10.1080/10683160601060374

Menghini, M., Ducro, C., & Pham, T. H. (2005). Évaluation de la récidive auprès d'agresseurs sexuels issus d'un hôpital sécuritaire en Belgique francophone. *Psychiatrie et Violence, 51*(1), 39-48.

Moens, I., & Pauwelyn, L. (2012). *Geen opsluiting, maar sleutels tot re-integratie. Voorstellen voor een gecoördineerd zorgtraject voor geïnterneerden*. Brussel: Zorgnet Vlaanderen.

Morgan, R. D., Flora, D. B., Kroner, D. G., Mills, J. F., Varghese, F., & Steffan, J. S. (2012). Treating offenders with mental illness: A research synthesis. *Law and Human Behavior, 36*(1), 37-50. doi: 10.1037/h0093964

Müller-Isberner, R., Freese, R., Jockel, D., & Gonzalez Cabeza, S. (2000). Forensic psychiatric assessment and treatment in Germany. Legal framework, recent developments, and current practice. *International Journal of Law and Psychiatry, 23*(5-6), 467-480.

Nilsson, T., Wallinius, M., Gustavson, C., Anckarsäter, H., & Kerekes, N. (2011). Violent recidivism: A long-time follow-up study of mentally disordered offenders. *PLoS ONE, 6*(10), e25768. doi: 10.1371/journal.pone.0025768

Nowak, M. J., & Nugter, M. A. (2014). 'Ik zit mijn tijd wel uit...' II. Een replicatieonderzoek naar recidiven bij patiënten met een strafrechtelijke machtiging ['I'll serve my time...' II. A replication study of recidivism in patients with a one-year hospital order]. *Tijdschrift voor Psychiatrie, 56*(4), 237-246.

Perlin, M. L. (2014). "The ladder of the law has no top and no bottom": how therapeutic jurisprudence can give life to international human rights. *International Journal of Law and Psychiatry, 37*(6), 535-542. doi: 10.1016/j.ijlp.2014.02.026

Pham, T. H., & Ducro, C. (2008). Évaluation du risque de récidive en Belgique francophone: Données préliminaires d'analyse factorielle de la "Sex Offender Recidivism Appraisal Guide" (SORAG) et de la Statique-99 [Risk assessment in social defence: Preliminary factorial analysis of the SORAG "Sex Offender Recidivism Appraisal Guide" (SORAG) and the Static-99. *Annales Médico-Psychologiques, 166*(7), 575-579. doi: 10.1016/j.amp.2008.06.001

Pham, T. H., Ducro, C., Marghem, B., & Réveillère, C. (2005). Evaluation du risque de récidive au sein d'une population de délinquants incarcérés ou internés en Belgique francophone [Prediction of recidivism among prison inmates and forensic patients in Belgium]. *Annales Médico-Psychologiques, 163*, 878-881.

Robert, L., & Maes, E. (2012). *Wederopsluiting na vrijlating uit de gevangenis*. Brussel: NICC.

Schmidt-Quernheim, F., & Seifert, D. (2014). Evaluation der ambulanten Nachsorge forensischer Patienten (§ 63 StGB) in Nordrhein-Westfalen [Evaluation of outpatient aftercare of forensic patients (§ 63 StGB) in North Rhine-Westphalia, Germany]. *Nervenartz, 85*(9), 1133-1143.

Schuringa, E., Spreen, M., & Bogaerts, S. (2014). Inter-rater and test-retest reliability, internal consistency, and factorial structure of the Instrument for Forensic Treatment Evaluation. *Journal of Forensic Psychology Practice, 14*(2), 127-144. doi: 10.1080/15228932.2014.897536

Seifert, D., & Moller-Mussavi, S. (2005). Preliminary recidivism rates of the Essener prognosis study. *Fortschritte der Neurologie-Psychiatrie, 73*(1), 16-22. doi: 10.1055/s-2004-830062

Skipworth, J., Brinded, P., Chaplow, D., & Frampton, C. (2006). Insanity acquittee outcomes in New Zealand. *Australian and New Zealand Journal of Psychiatry, 40*(11-12), 1003-1009. doi: 10.1111/j.1440-1614.2006.01924.x

Stassen, W., Habets, P., Mertens, A., Laender, J. De, & Jeandarme, I. (2014). The InReach project: From penitentiary to forensic hospital. *Therapeutic Communities: The International Journal of Therapeutic Communities 35*(3), 119-126.

Tabita, B., Santi, M. G. de, & Kjellin, L. (2012). Criminal recidivism and mortality among patients discharged from a forensic medium secure hospital. *Nordic Journal of Psychiatry, 66*(4), 283-289. doi: 10.3109/08039488.2011.644578

Vandevelde, S., Soyez, V., Beken, T. Vander, Smet, S. De, Boers, A., & Broekaert, E. (2011). Mentally ill offenders in prison: The Belgian case. *International Journal of Law and Psychiatry, 34*(1), 71-78. doi: 10.1016/j.ijlp.2010.11.011

Vertommen, H., Verheul, R., Ruiter, C. de, & Hildebrand, M. (2002). *Handleiding bij de herziene versie van Hare's Psychopathie Checklist*. Lisse: Swets Test Publishers.

Vitacco, M. J., Rybroek, G. J. Van, Erickson, S. K., Rogstad, J. E., Tripp, A., Harris, L., & Rodney, M. (2008). Developing services for insanity acquittees conditionally released into the commmunity: Maximizing success and minimizing recidivism. *Psychological Services, 5*(2), 118-125. doi: 10.1037/1541-1559.5.2.118

Vitacco, M. J., Vauter, R., Erickson, S. K., & Ragatz, L. (2014). Evaluating conditional release in not guilty by reason of insanity acquittees: A prospective follow-up study in Virginia. *Law and Human Behavior, 38*(4), 346-356. doi: 10.1037/lhb0000071

Ward, T., & Stewart, C. (2003). Criminogenic needs and human needs: A theoretical model. *Psychology, Crime & Law, 9*(2), 125-143. doi: 10.1080/1068316031000116247

Wartna, B. S. J. (2009). *In de oude fout. Over het meten van recidive en het vaststellen van het succes van strafrechtelijke interventies*. Den Haag: WODC.

Wechsler, D. (2005). *WAIS-III Nederlandstalige bewerking.Technische handleiding*. Lisse: Swets & Zeitlinger.

Wiederanders, M. R. (1992). Recidivism of disordered offenders who were conditionally vs unconditionally released. *Behavioral Sciences & the Law 10*, 141-148.

Wiederanders, M. R., Bromley, D. L., & Choate, P. A. (1997). Forensic conditional release programs and outcomes in three states. *International Journal of Law and Psychiatry, 20*(2), 249-257.

Yiend, J., Chambers, J. C., Burns, T., Doll, H., Fazel, S., Kaur, A., . . . Fitzpatrick, R. (2011). Outcome measurement in forensic mental health research: An evaluation. *Psychology, Crime & Law, 17*(3), 277-292. doi: 10.1080/10683160903203938

Relapse after treatment of forensic psychiatric patients under the authority of the Commission of the Protection of Society Ghent

Panopticon, 36(3), 227-247

11

Jeandarme, I., Pouls, C., Wittouck, C., Vander Laenen, F., Ampe, M., Verelst, R., Degrauwe, S., Oei, T. I., & Bogaerts, S.

ABSTRACT

Recidivism rates for forensic medium security patients (internees) in Flanders are lacking. The current article discusses recidivism rates in a medium security population under the authority of the Commission of the Protection of Society (CPS) Ghent. The research consisted of two studies. In the first study, recidivism was based on official convictions; in the second study a second outcome measure was taken into account (incidents). The base rate in the first study was 13% for general and 7% for violent recidivism. This increased to 40% in the second study, indicating that incidents are often not brought to court. Habitual offenders recidivated more than occasional offenders.

INTRODUCTION

Flemish internees reside under the jurisdiction of a Commission of the Protection of Society (CPS) and stay in the psychiatric wards of a prison or in penal institutions for Social Defense (Merksplas, Turnhout or Bruges). They are treated in a medium security unit (MSU) or within general mental health care (Vander Laenen & the Cauwer, 2011). The current study focuses on the three MSUs that were established in 2011 within the Public Psychiatric Care Center Rekem, the psychiatric center Sint-Jan-Baptist Zelzate, and the university psychiatric center St. Kamillus in Bierbeek. The term *medium security* refers to a medium security risk[1] to themselves and to society, and to an average risk of relapse (Moens & Pauwelyn, 2012). Forensic MSU treatments aim to reduce the risk of relapse to an acceptable level so that inmates can be transferred to a less secure environment. The MSUs in Flanders focus on Dutch-speaking internees with some degree of motivation and learning ability and, with respect to psychopathology, preferably with a psychotic and/or a personality disorder. Exclusion criteria are sexual or addiction problems and/or a high degree of psychopathy (De Smedt, Mariën & Vermeiren, 2008). Since 2006, only MSU Zelzate has treated female internees. Since these MSUs were first started in 2001, the number of forensic beds has expanded and was supplemented with beds in forensic psychiatric care units and forensic sheltered homes. Despite the development of this forensic circuit, the MSUs still do not have a structural funding. As of January 1, 2007, forensic beds reside under the Federal Public Service (FPS) of the Ministry of Health, which allocates structural funding for 40 hospital beds per site. Additional funding is provided for the extra security requirements of the hospital beds, forensic psychiatric care, and integrated living, and is governed by a yearly renewable agreement between the promoters and the FPS of the Ministry of Health (Trieste, 2009).

Despite the government's efforts since 2001, the treatment options for internees remain inadequate. In August 2012, there were still 699 detainees residing in a Flemish prison due to lack of referral possibilities (Moens & Pauwelyn, 2012). The effectiveness of forensic treatment is ideally measured over a wide range of indicators, including clinical and humanitarian outcomes such as quality of life. One of the most important indicators is criminal recidivism (Wartna, 2009; Yiend et al., 2011). Nevertheless, studying recidivism in a forensic context appears to be

associated with many methodological difficulties. Studies in the forensic field use divergent outcomes measures to operationalize recidivism, which are based on different stages in the criminal process (e.g., arrest, conviction, or re-incarceration in prison) and various methods (e.g., self-reports and standardized instruments). In addition, the follow-up periods and the operationalization of an offense vary (Chambers et al., 2009). Generally, conviction data are considered one of the most reliable sources of recidivism, even though they underestimate the number of crimes actually committed (Wartna, 2009). In the current study, a distinction is made between general and violent recidivism. General recidivism refers to any crimes, regardless of the type of offense. Violent recidivism refers to relapses into criminal acts of a violent nature, such as murder, assault and battery, rape, threat, or robbery with threat.

Recidivism rates for internees in Belgium

National recidivism rates of internees are not available in Belgium. However, one national survey was carried out among prisoners (including internees), which calculated the number of re-incarcerations after release from prison. The authors found a high rate of re-incarceration (62.3%) after a follow-up period of 5.7 to 8.7 years (Robert & Maes, 2012). Although this study did not include a *pure* measure of recidivism[2], the high re-incarceration figure is food for thought.

In three Walloon studies in (partially overlapping) populations of interned sexual offenders, reconviction rates ranged from 33.1% to 34.1% for general recidivism, from 5.2% to 17.1% for violent recidivism, and from 25.0% to 26.1% for sexual recidivism, each over a time span of about 2.5 to 4.2 years (Ducro & Pham, 2006; Menghini, Ducro, & Pham, 2005; Pham & Ducro, 2008). A study of 110 interned (sexual) violent offenders showed 21.0% general and 9.1% violent recidivism over 3.4 years (Pham, Ducro, Marghem, and Réveillère, 2005). In Flanders, no recidivism data have been published (Moens & Pauwelyn, 2012).

International recidivism data in forensic psychiatric populations

A comparative study in 2006 showed that 19 of the 41 European countries - including Belgium - had no national recidivism data (Wartna & Nijssen, 2006). In populations found (partially) irresponsible, recidivism data are only available on subpopulations[3]. In a Dutch study in patients found completely irresponsible, 43.5% relapsed one year after judicial placement (Nowak &

Nugter, 2014)[4]. In Sweden, general recidivism rates ranged from 8.0% to 63.0% with increasing follow-up periods (6.1 and 8.2 years, respectively) and 7.0%, 23.8%, and 46.0% violent recidivism (follow-up periods 3.6, 6.1, and 8.2 years, respectively; Lund, Hofvander, Forsman, Anckarsater, & Nilsson, 2013; Nilsson, Wallinius, Gustavson, Anckarsäter, and Kerekes, 2011; Tabitha, the Santi & Kjellin, 2012). In Germany, 6.6% relapsed with a violent crime after 2.5 years (Müller-Isberner, Freese, Jockel, Cabeza & Gonzalez, 2000). In medium security populations in England, after two years the rates for general recidivism were 10.4% and 15.0%, and for violent recidivism it was 6.0% (Coid, Hickey, Kahtan, Zhang, & Yang, 2007; Edwards, Steed, & Murray, 2002; Friendship, McClintock, Rutter, & Maden, 1999; Maden, Scott Burnett, Lewis, & Skapinakis, 2004). After a follow-up period of 6 years, these figures increased to 30.4% and 34.0% general recidivism and 18.0% violent recidivism (Coid et al., 2007; Friendship et al., 1999), and after 9.4 years the rates increased further to 47.8% general recidivism and 14.4% serious violent recidivism (Davies, Clarke, Hollin, & Duggan, 2007).

In Canada, 19.6% recidivated after three years, 22.9% after four years, and 36.0% after ten years with a violent crime (Rice, Harris, & Long, 2013; M. Rice, personal communication, February 27, 2014). An older Canadian study found lower relapse rates (7.3% general and 1.8% violent recidivism after 6.7 years; Luettgen, Chrapko, & Reddon, 1998). Data from the United States have been varied. In the state of Wisconsin, 7.1% general recidivism was found and 3.7% violent recidivism after 2.9 years (Vitacco et al., 2008). In the city of Saint Louis, 14.5% relapsed after having successfully completed treatment and 38.0% relapsed after dropping out from treatment over a follow-up period of one year (Dirks-Linhorst & Linhorst, 2010). In the state of California, overall recidivism rates of 50.8% have been reported after two years (Lee, 2003). In the city of Los Angeles (i.e., within the state of California), after 1.9 years, 10.1% were re-arrested (Bertman-Pate et al., 2004). Bonta, Blais and Wilson (2014) reported, in their meta-analysis of mainly North American studies, 39.0% general and 23.0% violent recidivism after an average period of 4.9 years. In New Zealand, 15.0% general recidivism and 5.8% violent recidivism was found after two years and 40.0% general recidivism after 10 years (Skipworth, Brinded, Chaplow, & Frampton, 2006). In Japan, finally, after a median follow-up period of 10.8 years, 10.6% was re-arrested or adjudicated because of a violent offense (Yoshikawa et al., 2007).

The abovementioned studies show that reconviction rates in and outside of Europe vary greatly, more specifically from 7.1% to 63.0% for general recidivism and from 1.8% to 46.0% for violent recidivism. Comparisons between studies are hampered by methodological differences. Major differences include the length of follow-up periods and the operationalization of outcome measures. For example, new convictions were used in more than half of the studies discussed (52.6%), whereas other studies also include re-arrests (26.3%) or even broader outcome measures such as re-admission to a forensic hospital because of new crimes (21.1%). The context in which recidivism takes place also influences recidivism rates. Some studies include re-admission periods and/or detention periods (among others, Friendship et al., 1999; Lund et al., 2013), whereas other studies exclude all periods of residential stay (Coid et al., 2007). Morgan et al. (2012) stressed the obvious shortage of good studies, which has not allowed for any conclusions on treatment effects, although there was a positive association with some behavioral characteristics[5]. In another meta-analysis, an overall reduction in criminal interventions was found, such as a decrease in the length of detention and the number of arrests after treatment. Also, there was a positive trend for the number of new convictions, whereas conditional release violations showed a negative correlation with treatment (Martin, Dorken, Wamboldt, & Wootten, 2012). In Sweden, after a follow-up period of 13 to 20 years, Lund et al. (2013) found no difference between offenders who received a prison sentence versus offenders who received an alternative sanction after a forensic examination, which contrasted earlier findings over a shorter follow-up period (Nilsson et al., 2011).

Predictors of violent recidivism in forensic psychiatric populations

Predictors of violent recidivism in forensic psychiatric patients are similar to those in mainstream offender populations (Bonta et al., 2014). Among those risk factors, having a personality disorder (Coid et al., 2007) or a substance misuse disorder, especially alcohol abuse (Bonta et al, 2014; Spreen, Fire, Ter Horst & Bogaerts, 2014) is associated with violent recidivism. Of the eight risk factors that have been repeatedly validated in the literature (Central Eight), in particular an antisocial personality pattern (and more specifically psychopathy), procriminal attitudes, and a history of antisocial behavior were the best predictors for violent recidivism (Bonta et al., 2014). In addition, Wartna et al. (2005) found a correlation between violent recidivism and male sex,

habitual offenders, young age at discharge, Dutch origin, type of index offense, lack of trial leave, and absconding. The number of antecedents and (young) age at the time of first trial proved to be an explanatory factor (Coid et al., 2007; Nilsson et al., 2011). On the other hand, severe mental illness such as psychosis appears to be a rather weak or negative predictor of violent recidivism. Intellectual disabilities are weakly related to crime, although intellectually disabled patients have been over-represented in forensic psychiatric units (Cullen, Gendreau, Jarjoura, & Wright, 1997; Lunsky et al., 2011; Maes, Goethals, & Verlinden, 2009).

There are currently more than 100 tools available to predict the risk of (violent) recidivism, among which the Psychopathy Checklist-Revised (PCL-R; Hare, 2003) and the Historical Clinical Risk Management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997) are the most frequently used in practice (Hurducas, Singh, de Ruiter & Petrila, 2014; Khiroya, Weaver, & Maden, 2009; Singh et al., 2014). These instruments also have predictive validity in forensic psychiatric populations (Dolan & Khawaja, 2004; Gray, Taylor, & Snowden, 2008; Leistico, Salekin, DeCoster, & Rogers, 2008).

Clinically, non-compliance[6] is often regarded as a marker for relapse, but so far has been little examined in forensic populations. Non-compliance did appear to be associated with a higher risk of readmissions (Bertman-Pate et al., 2004) and with violent incidents during treatment (Jeandarme et al., 2015) in forensic psychiatric patients. Non-compliance has also been associated with repeat partner violence (Kindness et al., 2009).

Victims of violence by forensic psychiatric patients

Studies show that victims of violence by psychiatric patients are usually acquaintances of the offender, such as family members, friends or acquaintances, or professionals such as social workers or the police (Joyal, Dubreucq, Gendron, & Millaud, 2007; Nordstrom & Kullgren, 2003; Steadman, Mulvey, Monahan et al., 1998). Compared with the general population, psychiatric patients themselves have a higher chance of being a victim (e.g., Choe, Teplin, and Abram, 2008).

The current research

The current study provides a descriptive overview of the reconviction rates within a subpopulation of internees after discharge from a MSU (*Study 1*). In addition, a number of more

specific research questions were formulated. First, it was expected that, due to dark number effects, recidivism rates that include incidents would be significantly higher compared to recidivism rates based only on official criminal records (*Study 2*). The second hypothesis was that, following the international literature, acquaintances would more likely become victims of violent offenses than strangers. Because MSUs focus primarily on the treatment and prevention of aggressive behavior, internees with and without (sexual) violent offenses were compared along a number of demographic, clinical, offense, and treatment-related variables. Hereby, as a third hypothesis, it was expected that historical factors such as age at first sentence or the number of prior sentences would better predict violent recidivism than clinical factors such as presence of a major psychiatric disorders. Without a control group – internees or convicts with a mental disorder who were released without treatment – no definitive statements can yet be made regarding a possible treatment effect. The assessment of the effectiveness of treatment is no easy task, given the multitude of factors that (may) influence recidivism (black box). Nevertheless, an attempt was made to map the evolution of criminal behavior before and after treatment.

METHOD

Procedure

Between 2001 and 2010, 203 internees under the jurisdiction of CPS Ghent were treated in a MSU. Nine were not discharged at the end of the study and two passed away during treatment. Of the final study population (N = 192), 64.1% were treated in Zelzate (n = 123), 20.3% in Bierbeek (n = 39), and 15.6% in Rekem (n = 30). Information on the characteristics of the internees was collected through CPS files and MSU hospital records. By making use of secondary data, it was not always possible to ascertain whether the information was collected in a valid and reliable manner (Lievens, 2001). Wherever possible, missing data were added, but in particular the PCL-R and HCR-20 were not always assessed. On the other hand, one strength of the study is that it provides a picture of real world practices.

In Study 1, conviction rates (hereinafter *recidivism*) were collected during and after expiry of the index internment through the extraction of Central Criminal Records of the Federal Public Service of the Ministry of Justice on February 16, 2012. These records included not only

convictions but also community services, internment measures, fines, and probation. The date of the offense was retrieved from the prosecutors. Study 2 examined whether crime-related incidents reported to the CPS and which occurred during the index internment period[7] (hereinafter *incidents*), differed from the official recidivism data.

The criminal response to the incident reports was investigated through 1) the decisions of the CPS regarding condition violations, 2) the detention files from the prison administrative registration system (*SIDIS*), and 3) the Central Criminal Records, combined with information from the public prosecutor.

Outcome measures

In Study 1, recidivism was defined as any new conviction and/or internment that was pronounced after the end of the (first) treatment[8]. Violent recidivism related to a new conviction and/or internment for a (sexual) violence offense. Criminal Records were chosen as the outcome measure for recidivism because they can be regarded as a more reliable source for recidivism data, compared to self-report and police arrests (Wartna, 2009). However, this presupposes a reliable and up-to-date registration system (Lee, 2003). For example, the current study showed that in 33.5% of cases, prior sentences retrieved from the Criminal Records differed when comparing older to more recent extracts. Therefore, if the extracted records differed then all sentences were included. In addition, there is a time lag between the offense, the actual sentencing, and registering the sentencing date into the Central Criminal Records[9]. Furthermore, adjudicated offenses reflect an underreporting of actual re-arrest or detection rates by the police. Some offenses are not detected or not prosecuted[10].

To obtain a more complete picture of recidivism, a second outcome measure was consulted, namely crime-related incidents, reported to CPS Ghent. The choice for this second outcome measure makes the research design stronger. Yet it remains inadequate in response to the dark number problem. In hospital and prison settings, there is no uniform method of reporting to the CPS or the judicial assistant. Also, after termination of the internment measure, individuals can disappear from the CPS radar. In Study 2, a crime-related incident was defined as an incident coded under offending categories of the Belgian penal code, whether or not they led to further prosecution or sentencing. These incidents were classified into five categories, as

defined in Table 1. In cases of multiple incidents per report, the classification was based on the most serious incident (Brand, 2005).

Table 1. Definition of crime-related incidents

(Sexual) violent offense	Homicide and manslaughter, arson with danger to persons, property crime with violence, sexually violent offense against minors or adults, stalking, and other violent crimes
Property offense	Property offense without violence (e.g., theft, fraud) and arson without endangering persons
Offenses against drug laws (hereafter drug offense)	Possession, use and distribution of illegal drugs and medication which was not prescribed
Other offense	Traffic violations and other crimes such as vandalism or disturbing the public order
Non-violent sexual offense	Hands-off offense (e.g., exhibitionism)

For Studies 1 and 2, the prevalence of recidivists was calculated, together with the frequency and nature of the relapse. It was also examined whether direct victims were involved in (verbal) violent crimes and the nature of the relationship between the internee and victim.

In Study 1, for each participant, a comparison was made between the annualized rate of sentences pre- and post-treatment to assess the evolution in criminal behavior according to the formulas (Gilleir, Easton, & Ponsaers, 2010):

$$\text{Density sentences pre-treatment} = \frac{\text{number of sentences before admission}}{\text{Time period (18 years until first admission)}} \times 365$$

$$\text{Density sentences post-treatment} = \frac{\text{number of sentences after admission}}{\text{Time period (first discharge until 31/12/2010 or death)}} \times 365$$

A more stringent study design is necessary in order to examine treatment effects. A control group – internees or convicts with mental disorder exempted without treatment – was unfortunately not available (Sherman et al., 1998).

Finally, in Study 2, internees with (n = 43) and without (n = 102) (sexual) violent offenses (incident and/or recidivism) were compared on a number of demographic, clinical, offense, and treatment-related variables. Only internees with a minimum follow-up period of two years were included in these analyses.

Demarcation follow-up period

The antecedent period ran from the participant's eighteenth birthday to the date of first MSU admission. The entire follow-up period started at the first discharge date from MSU and ran until the end of the census date of the study (31/12/2010) or death. This period was divided into: (1) a period in which the index internment was still in force (i.e., beginning at discharge date until the internment expired or death) and (2) a period after unconditional release (i.e., beginning the day after final expiration until census date). Periods of readmission at a medium security department were not included (and deducted from) the follow-up period[11]. The mean follow-up period during internment index (*N* = 192) was 3.2 years (1,156.18 days, *SD* = 719.00 days; range = 26–3,150 days) and after unconditional release (*n* = 44) was 3.0 years (1,080.09 days, *SD* = 598.44 days; range = 44–2,279). The entire follow-up period averaged 3.8 years (1,403.76 days, *SD* = 795.56 days; range = 26–3,165).

Participants

The vast majority of the internees were male (92.2%; *n* = 177).[12] In three cases, nationality could not be traced; in other cases (*n* = 189), the majority were Belgian nationals (95.2%, *n* = 180). The average age at time of the first admission was 35.7 years (*SD* = 10.27; range = 18.77–73.38). The mean hospital stay (including any re-admissions) was 1.7 years (624.98 days; *SD* = 471.72 days; range = 8–2,729 days). One third of the population (36.5%; *n* = 70) was referred back to prison or another medium security department at first discharge. Of the remaining internees (*n* = 122), almost two-thirds (55.7%, *n* = 68) continued treatment in a less secured forensic unit and more than a third (44.3%; *n* = 54) continued with regular psychiatric care. Nearly one-fourth of the population (30.7%; *n* = 59) absconded during MSU treatment. A fourth of the population (24.0%; *n* = 46) was non-compliant (e.g., refusal of therapy or medication, drink of alcohol).

The most common DSM-IV-TR diagnoses (American Psychiatric Association, 2000) were personality disorders (65.6%; *n* = 126), psychotic disorders (45.3%; *n* = 87), and substance misuse problems (45.3%; *n* = 87). The diagnoses were clustered according to the classification system used by Monahan et al. (2001): a) major mental disorders on Axis I disorders (MMD) (38.0%, *n* = 73); b) major mental disorders in combination with substance misuse (MMD-SUD; 17.7%; *n* = 34), including directly-related complications; and c) other disorders (OTH; 44.3%, *n* = 85), which

included personality disorders, learning disabilities, addiction problems, or a combination of these. A substantial proportion of the population (42.7%; *n* = 82) had no available Wechsler Adult Intelligence Scale-III (WAIS-III; Wechsler, 2000) score. The average WAIS-III IQ score for the remaining internees (57.3%, *n* = 110) was 78.0 (*SD* = 17.41, range = 48–135).

In 132 internees (68.8%), a PCL-R score[13] was not available. For the rest (31.3%, *n* = 60), the average PCL-R score was 20.0 (*SD* = 7.06; range = 8 to 32.6) and almost one-third (31.7%; *n* = 19) scored above the cut-off score of 25. For 143 internees (74.5%), no HCR-20 was available[14]. Of the remaining internees (25.5%, *n* = 49), nearly one-third (22.5%; *n* = 11) were classified as high risk on the basis of the HCR-20.

A minority (15.1%; *n* = 29) were first offenders at the index internment measure. An internee was considered a habitual offender when he had three or more criminal sentences as an adult within the five years prior to the index offense (Wartna & Tollenaar, 2004). In 16.7% of the population (*n* = 32), due to age it was impossible to identify a five year period prior to index internment (i.e., internees between 18 and 22). In the remaining internees (*n* = 160), 48.8% (*n* = 78) were considered a habitual offender. The average age at first verdict was 25.0 years (*SD* = 9.77, range = 9.80–66.54). The (most severe) index offense that gave rise to the decision to be included in an MSU was usually either a (sexual) violence (74.5%) or property crime (22.9%). The majority of the population (94.3%, *n* = 181) had committed previous (sexual) violent crimes.[15]

Data analysis

Descriptive statistics and analyses were conducted in SAS version 9.3, R, and SPSS version 22. Continuous variables were expressed as means, standard deviations, and ranges. Categorical variables were reflected via percentages in each category. Background density was compared to recidivism density using a paired t-test. The association between relapses in violent crimes and each of the socio-demographic, offense, and treatment-related characteristics were separately examined by means of the Pearson chi-square and Fisher exact test. The corresponding relative risk (RR) was calculated with the 95% confidence interval. The difference in mean age at time of first sentence and at first admission between violent and non-violent participants was tested using an unpaired t-test (after transformation of the data where the data was not distributed normally). If no proper transformation of the data was required, the means of the two groups

were compared using the Mann Whitney U test. The difference in the average number of antecedents was tested with a negative binomial regression. Characteristics that had a significant association with violent crime relapse were then entered into multiple logistic regression. All tests were two-sided with significance level of 5%.

RESULTS AND DISCUSSION

Results Study 1: Recidivism

The following results refer to the period in which the index internment was still in effect as well as the period after unconditional release.

Descriptive data

Prevalence. Table 2 shows the prevalence of recidivists. About one-tenth (13.0%) of all internees received a new conviction and/or internment measure. Only 14 internees in the total sample (7.3%) relapsed with a violent offense. Compared to international figures, the number of recidivists – 8.3% excluding fines – was low (e.g., Coid et al., 2007; Friendship et al., 1999; Lund et al., 2013; Nilsson et al., 2011; Yoshikawa et al., 2007). The current follow-up period excluded periods of MSU readmissions, in line with the recommendation of Skipworth et al. (2006). The low recidivism rates may be due to different reasons. For example, the mean follow-up period in this study was generally smaller than is described in the literature. Also the treatment and monitoring procedures (infra) can differ on several topics between countries. Also, the use of criminal records potentially led to underreporting. In almost half of the international studies, other outcomes were used, such as a new arrest (e.g., Dirks-Linhorst & Linhorst, 2012; Yoshikawa et al., 2007). Some of the Walloon studies relied on new convictions known to the CPSs. In the present study, both CPS files and recent Criminal Records extracts were analyzed, which revealed that more than half of the recidivism offenses (63.4%, n = 26/41)[16] were not known to the CPS.

Another striking finding was that as many or even more internees incurred a new conviction (8.9%, n = 17; excluding fines 4.2%, n = 8) as received a new internment measure (4.7%; n = 7)[17]. Other studies also reported that after forensic treatment, patients were later on convicted to an imprisonment sentence (Friendship et al., 1999).

Table 2. Prevalence of recidivists *(N = 192)*

	Convicted		Interned		Total	
	n	%	*n*	%	*N*	%
General recidivism[18]	17	8.9	9	4.7	25	13.0
(Sexual) violent recidivism	8	4.2	7	3.7	14	7.3

Frequency and density recidivists. The majority of the recidivists recidivated on one occasion (84%; *n* = 21/25). On average, a recidivist received 1.4 new sentences (*SD* = 1.12, range = 1–6)[18]. This seems to be lower than the rate found in studies with a similar follow-up period (average 2.0 to 3.1, e.g., Edwards et al., 2002; Lee, 2003). The average pre-treatment offense density (0.45) was significantly higher than the post-treatment offense density (0.03; $p < .01$). The internees committed significantly fewer criminal offenses on an annual basis after a (first) MSU admittance than before.

Nature of recidivism (Table 3). More than half (56%; *n* = 14/25) of the recidivists recidivated with a (sexual) violence offense. In addition, one individual committed a property crime (4%; *n* = 1) and one committed a drug offense (4%; *n* = 1). A third (36%; *n* = 9) relapsed with another offense, mainly traffic offenses. In the current study, the concept of violence has been defined broadly. Not all international studies give a definition, but presumably verbal violence would not always be included in all studies, whereas most studies included sexual and property violence. Judging from the nature of the recidivism offenses, it can be concluded that they can be regarded as less serious (i.e., no homicides, 6.3% (*n* = 12/192) nonverbal violence including one sexual violence offense). There was, for example, only one conviction that led to imprisonment for more than six months. As with previous research findings, relapses after forensic treatment proved to be less severe compared with criminal offenses before treatment (Lee, 2003).

Context of relapse at the time of offense (Table 3). Recidivism mainly occurred after the index internment (22.7%, *n* = 10/44 after unconditional release vs. 7.8%, *n* = 15/192 during index internment). The degree of supervision or lack thereof appears to play an important role, which is confirmed in the literature (Bogaerts, Willems, Spreen, Schuringa, and Ter Horst, 2013; Dirks-

Linhorst & Linhorst, 2012). Another indication of this can be found in the fact that most of the patients relapsed during ambulatory supervision. Of the 15 internees who relapsed during the index internment, two-thirds (46.7%, *n* = 7) were under outpatient supervision in the community, a quarter (26.7%, *n* = 4) were detained, and a quarter (26.7%) stayed in a forensic psychiatric care center (*n* = 2) or a general psychiatric hospital (*n* = 2).

Table 3. Context and nature of recidivism using the most serious offense

	Recidivism during index internment				Recidivism after index internment	Total
	Society with supervision	Prison	ForCare	GPH	Society without supervision	
General recidivism	7	4	2	2	10	25
(Sexual) violence offense	4	3	1	1	5	14
Homicide/manslaughter	-	-	-	-	-	-
Verbal violence	-	1	-	1	-	2
Sexual violence	1	-	-	-	-	1
Property offense with violence	1	-	-	-	-	1
Other violence	2	2	1	-	5	10
Property Crime	-	-	1	-	-	1
Drug offenses	1	-	-	-	-	1
Traffic offense	1	1	-	1	5	8
Other offense	1	-	-	-	-	1

Note. ForCare = Forensic psychiatric Care center; GPH = General Psychiatric Hospital.

Victim information

In total there were 32 victims, of which the majority were men (65.6%; *n* = 21) and adults (90.6%; *n* = 29). In one case, the relationship with the victim could not be traced. In 42% of cases (*n* = 13/31) the victim was a stranger; in other cases (58.1%, *n* = 18/31) there was a personal or professional relationship with the victim[19]. These findings are in line with the literature (Nordstrom & Kullgren, 2003; Yoshikawa et al., 2007).

Results Study 2: Incidents

The following results refer only to the period in which the index internment was still in effect.

Descriptive data

Prevalence (Table 4). After MSU treatment, more than one-third of the total sample (39.1%, *n* = 75) was involved in at least one general incident, and 26.0% (*n* = 50) were involved in a (sexual) violence incident. A minority (7.8%; *n* = 15) was convicted for a general (7.8%; *n* = 15) or a violent offense (4.7%; *n* = 9). In two-thirds of the offenders (66.8%, *n* = 10/15), a verdict was pronounced following an incident report. In a third of the offenders (33.3%, *n* = 5/15), the recidivism offense was not recorded in the CPS file.

As expected, there was an increase in the number of internees that relapsed when incident reports were added to the recidivism data. In line with the literature, the recidivism rate was five times higher when a broad measure of recidivism was used (Falshaw, Bates, Patel, Corbett & Friendship, 2003). Significantly, more internees were involved in a general incident than in recidivism (McNemar, $p < .01$). The same pattern was evident in the violent incidents compared to violent recidivism (McNemar, $p < .01$). The recidivism figures that were based on the Central Criminal Records (7.8%) increased substantially when criminal incidents from the CPS files were included (40.1%), suggesting that these incidents were not prosecuted. On the other hand, in half of the internees with an incident and/or recidivism (44.2%, *n* = 34/77), conditional release was revoked at least once, which often resulted in a long detention period.[20] By contrast, only a minority of the incident perpetrators (13.3%, *n* = 10/75) was adjudicated after an incident had been reported to the CPS.

Table 4. Prevalence of incident perpetrators and recidivists within the total population (*N* = 192)

	Incident		Recidivism		Total	
	n	%	*n*	%	*N*	%
General offense	75	39.1	15	7.8	77	40.1
(Sexual) violence	50	26.0	9	4.7	50	26.0

Frequency. About half of the incident perpetrators were involved in only one incident (48.0%, *n* = 36/75). A third (28.0%, *n* = 21/75) were involved in two or three incident reports and 24% (*n* = 18/75) were involved in four or more incident reports during the incident follow-up period. On average, an incident perpetrator was involved in 2.7 (*SD* = 2.70, range = 1–13) incidents.

On the other hand, most recidivists (86.7%; *n* = 13/15) only received one new sentence.

The combined relapse figures showed an increase in the number of offenders with multiple offenses (50.7%; *n* = 39/77).

Nature of the incidents (Table 5). Two-thirds of the incident perpetrators (65.3%, *n* = 49/75) relapsed in violent offenses, 16.0% (*n* = 12/75) in drug offenses, 14.7% in property offenses (*n* = 11/75), and 2.7% (*n* = 2/75) in non-violent sexual offenses. These percentages in recidivists (*n* = 15) were lower: 60.0%, 6.7%, 6.7%, and 0.0%, respectively.

Compared with the recidivisms, the incidents showed a five percentage point increase in the number of internees with violent offenses, in particular with regard to verbal and other violence. In addition, the percentage of internees with property crimes and drug offenses increased eight and ten percentage points, respectively.

Table 5. Prevalence of incident perpetrators and recidivists during the internment measure, based on most serious crime (*N* = 192)

	Incident		Recidivism		Total[21]	
	n	%	*n*	%	*N*	%
General offense	75	100	15	100	77	100
(Sexual) violence	49	65.3	9	60.0	49	63.7
Homicide/manslaughter	1	1.3	0	0.0	1	1.3
Verbal violence	14	18.7	2	13.3	13	16.9
Sexual violence	2	2.7	1	6.7	2	2.6
Property offense with violence	2	2.7	1	6.7	2	2.6
Other violence	30	40.0	5	33.3	31	40.3
Property offense	11	14.7	1	6.7	11	14.3
Drug offense	12	16.0	1	6.7	12	15.6
Sexual non-violence	2	2.7	0	0.0	2	2.6
Other offense	1	1.3	4	26.7	3	3.9

Context of the incidents.[22] Incident perpetrators remained, for the most part, in prison when the incident took place (29.3%, *n* = 22/75) or stayed in a forensic psychiatric care center (36.0%, *n* = 27/75).

When incidents and recidivism were combined (*n* = 77), nearly a fifth (19.5%, *n* = 15/77) occurred when the internee was under outpatient supervision in the community, 7.8% (*n* = 6/77) lived in a forensic sheltered home, 37.7% (*n* = 29/77) lived in a forensic psychiatric nursing, 13.0% (*n* = 10/77) were in a general psychiatric hospital, and 28.6% (*n* = 22/77) were in detention. There was a notable decline in the proportion of internees that relapsed while under outpatient supervision in the community. About half of the recidivists (46.7%, *n* = 7/15) recidivated while under outpatient supervision, whereas only 19.5% did so when incidents and recidivisms were combined.

Victim information

The incidents involved 145 victims, whereas only 18 victims were involved in the recidivisms. When incidents and recidivism were combined, the total number of victims rose to 147[23]. Nearly all of the victims were adults (97.9%). The ratio between male and female victims was approximately 1 to 2 (28 and 45), although information on the gender of the victim was often missing (*n* = 74). In most cases, there was a personal (38.1%) or professional (56.8%) relationship with the victim. Compared to the recidivism data, the combined (incident and recidivism) data showed a decline of about 15 percentage points with respect to the number of underage victims and an increase of 22 percentage points in the number of acquaintance victims.

Characteristics associated with violent behavior

No significant differences were observed in demographic and clinical characteristics between internees with and without a (sexual) violent re-offense (Table 6). However, a trend significant relationship ($p < .10$) was found for the PCL-R score. Also, it should be noted that the lack of significant results for the PCL-R and the HCR-20 can be attributed to the fact that in approximately two-thirds to three-quarters of the internees, no PCL-R score or HCR score could be found. Also, it is possible that the field validity of these instruments is poor. Research has shown that risk assessment schemes in general show good reliability and validity when scored by

Table 6. Socio-demographic, offense, and treatment characteristics in the total population, violent subpopulation and non-violent subpopulation

	Total population (*N* = 192)		Violent group (*n* = 43)		Non-violent group (*n* = 102)		*p*
Categorical variables	*n*	%	*n*	%	*N*	%	
Socio-demographic variables							
Male	117	92.2	41	95.4	97	95.1	1.00
Belgian nationality[a]	180	95.2	39	92.9	97	96.0	.42
Criminal records							
Habitual offender	78	48.8	23	67.7	35	40.7	.01*
Index internment offense							.50
Violence	143	74.5	29	67.4	78	76.5	
Property	44	22.9	13	30.2	22	21.6	
Drugs	3	1.6	1	2.3	1	1.0	
Sexual non-violence	1	0.5	0	0	1	1.0	
Other	1	0.5	0	0	0	0	
Treatment characteristics							
Absconding	59	30.7	20	46.5	26	25.5	.01*
Non-compliance	46	24.0	12	27.9	21	20.6	.39
Discharge to forensic unit	68	55.7	8	44.4	33	45.8	.92
Clinical characteristics							
PCL-R ≥ 25[b]	19	31.7	8	53.3	9	25.7	.06
HCR-20 high[c]	11	22.5	1	12.5	4	25.0	.63
Psychiatric diagnosis							.96
MMD	73	38.0	17	39.5	40	39.2	
MMD-SUD	34	17.7	8	18.6	21	20.6	
OTH	85	44.3	18	41.9	41	40.2	
Personality disorder	126	65.6	30	69.8	65	63.7	.48
Intellectual disabilities							.62
Normal	148	77.1	30	69.8	79	77.5	
Borderline intellectual functioning	12	6.3	4	9.3	7	6.9	
Intellectually disabled	32	16.7	9	20.9	16	15.7	
Continuous variables	*M*	*SD*	*M*	*SD*	*M*	*SD*	
Socio-demographic variables							
Age at first dismissal	37.0	10.21	34.7	9.71	37,6	9.74	.11
Criminal records							
Age at first sentence	25.0	9.77	22.7	9.03	25,85	9.11	.01*
Number of sentences	4.8	4.54	5.9	4.36	4,16	4.46	.08

Note. Column 1 shows the total population. The groups in column 3 (*n* = 43) and 4 (*n* = 102) relate only to persons with a minimum follow-up period of 2 years. HCR-20 = Historical, Clinical, Risk Management-20; PCL-R = Psychopathy Checklist-Revised; MMD = major mental disorders on Axis I; MMD-SUD = major mental disorder in combination with substance misuse; OTH = other disorders.
[a] 1.6% missing; [b] 68.8% missing; [c] 74.5% missing.
* $p < .05$.

trained researchers in a well-controlled research setting, but this is much less the case when they are scored by clinicians in the field. In the latter case, the results were less unambiguously positive, both in terms of inter-rater reliability and in terms of predictive validity (e.g., Murrie, Boccaccini, Caperton, and Rufino, 2012; Vojt, Thomson, & Marshall, 2013). The present study did not find a difference with respect to internee gender and presence of a personality disorder. Previous research found, for example, that men tend to recidivate more than women in violent crimes (Yoshikawa et al., 2007).

Regarding offense-related variables, the present study found – although limited in comparison with international studies – that past behavior predicts future behavior (Bonta et al., 2014; Coid et al., 2007; Nilsson et al., 2011; Wartna, el Harbachi, & van der Knaap, 2005). More specifically, significantly more habitual offenders recidivated with a violent offense, and they were 2.2 times more likely to relapse into a violent offense (p = .01; RR = 2.24; 95% CI RR [1.20; 4.17]). Also, younger age at first conviction was shown to predict relapse with violent behavior (p = .01, U = 1599; z = -2.57). Finally, the number of sentences was found to be a trend significant predictor.

Regarding treatment-related variables, a significant difference was observed between the groups for absconding. An internee who absconded from MSU treatment had a 1.8 times higher risk for violent recidivism (p = .02; RR = 1.87; 95% CI RR [1.15; 3.04]). This is in line with Dutch research (Wartna et al., 2005). There appeared to be no correlation with non-compliance, which can be explained by the fact that this variable was coded dichotomously (Kindness et al., 2009). Previous research in partner violence showed that it is important to define this variable quantitatively and to use multiple dimensions of non-compliance (Kindness et al., 2009).

When the characteristics that were associated with violent recidivism were combined into one model, only the variable habitual offender was independently associated with recidivism in violent crime (p = .04). The low base rate of recidivism – and consequently loss of statistical power – could be an explanation for the limited results that were found. In addition, other variables that were not investigated could potentially be of interest. For example, the presence of procriminal attitudes has been shown to be a significant predictor of violent recidivism (Bonta et al., 2014), a variable that could not be examined with the current data.

CONCLUSION

This is the first study to report recidivism data in Flemish internees. A minority of internees (13.0%) received a new conviction or internment measure after MSU treatment. Moreover, the recidivism was not particularly serious (i.e., no homicides, one sexual violent offense, 6.3% engaged in non-verbal violence). Violence was perpetrated primarily towards acquaintance victims. Although these data are encouraging for the MSUs, it cannot be concluded that the reduction in recidivism density can be attributed to treatment effects. Indeed, there are many factors that (may) influence whether or not a new offense will be committed (Sherman et al., 1998). Research examining the effects of judicial interventions in Flanders is virtually nonexistent. In a recent meta-analysis of recidivism in Dutch research, only one Flemish study was included (Robert, 2014).

Besides official recidivism, there are offenses that do not result in a new sentence. The present study showed that this occurred quite often. For example, only a minority of incident perpetrators (13.3%) were subsequently adjudicated. It appeared that these incidents were handled primarily through other mechanisms. Thus, almost half of the internees (44.2%), almost exclusively at the request of the MSUs, were locked up in prison due to the revocation of their conditional release. The subsequent detention periods were often long, and the risk of detention damage is real. Dropping out from or interrupting treatment is problematic and has been described in another article (Jeandarme et al., 2015). The rather significant base rate increase after the inclusion of incidents indicates the importance of the choice of the outcome measure. Future research should preferably make use of various outcome measures, including police arrest data. Another striking finding was the fact that, proportionally, three times more people relapsed after the end of the internment measure (concretely, 10 of the 44 internees with unconditional release recidivated), demonstrating the importance of supervision and risk management. It will be important to find a good balance between the need for control and the importance of offering a perspective in which unconditional release is a viable option (To, 2014).

A second important aspect of the study was the search for predictors of violent behavior. Contrary to the hypotheses, few differences were found between internees who relapsed with a

violent offense and internees who did not. Only habitual offenders were found to have a higher risk of relapse.

FOOTNOTES

[1] Security is a vital aspect in forensic care and consists of three aspects: 1) environmental or physical security, 2) procedural security, and 3) relational security (Kennedy, 2002). While the first two components emanate from the treatment provider's perspective, the third component is closely linked to the risk profile and the security needs of the patient (H. Vertommen, personal communication, August 13, 2013).The authors refer to the outcome measure as a by proxy recidivism measure. Whether these re-incarcerations are because of relapse into new crimes or merely due to violating the internment conditions, is not evident from this study.

[2] The authors refer to the outcome measure as a proxy for recidivism. Whether these re-incarcerations are because of relapse into new crimes or merely due to violating the internment conditions is not evident from this study.

[3] Inpatients under the Dutch Entrustment Act are not included in the (non-exhaustive) list because this population deviates too strongly from the interned population.

[4] Criminal disposition may be imposed for a maximum of one year in case of complete insanity.

[5] All treatments – including in prisons – for offenders with severe psychiatric disorders were enrolled in this study. Treatment effects were measured using control groups or pre- and post-measurements.

[6] Non-compliance means that patients do not comply with treatment agreements and regulations.

[7] Upon completion of the index internment, the CPS is no longer authorized; therefore no information was available with respect to incidents occurring after the internment measure.

[8] Convictions and/or measures pronounced in the period after the internment were only considered as recidivism when the crimes were effectively committed after discharge from treatment. Offense dates were retrieved from prosecutors. Offenses with sentencing dates after the end of the study that had offense dates occurring during the study period were also counted as recidivism.

[9] In the current study, the average time between offense date and sentencing date was 459.7 days (*SD* = 224.24; range = 71–953 days). In addition, data since 2006 are probably an underestimation due to a delay in the registration of the Central Criminal Records (Department for Criminal Policy, 2015).

[10] During the study period, each suspect had a unique number within one district, but no national number. As a result, it was not possible to collect information at the prosecutor's level.

[11] Relapse during MSU admission was described in Jeandarme et al. (2013).

[12] Since sex was not associated with the studied variables and because of the small number of female internees, analyses for men and women are not presented separately.

[13] Higher scores dimensionally indicate more psychopathic traits on the PCL-R. Scores range from 0 to 40. Categorically, in Europe, a cut-off score of 25 or more is regarded as indicative of psychopathy (Cooke & Michie, 1999).

[14] A high score on the HCR-20 indicates an increased risk of relapse into violence (Philipse et al., 2000).

[15] Index offense included.

[16] Besides recidivism offenses during the time at risk period (*n* = 35), there were recidivism offenses during MSU treatment (*n* = 3) and recidivism offenses with context unknown (*n* = 3).

[17] One internee had both a conviction and an internment measure.

[18] Event based there were 35 sentences, including fines (*n* = 15).

[19] Professionals include nurses and social assistants (*n* = 2), police officers (*n* = 4), and prison guards (*n* = 1).

[20] On average, incident perpetrators were re-incarcerated for 642.5 days.

[21] In some cases, the recidivism – as found in the Central Criminal Record – was also noted as an incident in the CPS files, and there was thus a double registration. In these cases, the recidivism took priority. Thus the overall prevalence is not always merely the sum of the incident and recidivism rates.

[22] Based on a more serious incident.

[23] In some cases, the victims of recidivism – as was found in the Central Criminal Record – were also listed as victims in the incidents in the CPS files, and there was thus a double registration. In these cases, the recidivism took priority. Thus the overall prevalence is not always merely the sum of the incident and recidivism rates.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Bertman-Pate, L. J., Burnett, D. M. R., Thompson, J. W., Calhoun, C. J., Deland, S., & Fryou, M. (2004). The New Orleans forensic aftercare clinic: A seven year review of hospital discharged and jail diverted clients. *Behavioral Sciences & the Law, 22*(1), 159-169. doi: 10.1002/bsl.575

Bogaerts, S., Willems, M., Spreen, M., Schuringa, E., & Horst, P. Ter (2013). *Van HKT-30 naar HKT-R*. Tilburg: Universiteit van Tilburg.

Bonta, J., Blais, J., & Wilson, H. A. (2014). A theoretically informed meta-analysis of the risk for general and violent recidivism for mentally disordered offenders. *Aggression and Violent Behavior, 19*(3), 278-287. doi: 10.1016/j.avb.2014.04.014

Brand, E. F. J. M. (2005). *Onderzoeksrapport PIJ-dossier 2003-C. Predictieve validiteit van de FPJ-lijst*. Den Haag: DJI.

Chambers, J. C., Yiend, J., Barrett, B., Burns, T., Doll, H., Fazel, S., . . . Fitzpatrick, R. (2009). Outcome measures used in forensic mental health research: A structured review. *Criminal Behaviour and Mental Health, 19*(1), 9-27. doi: 10.1002/cbm.724

Choe, J. Y., Teplin, L. A., & Abram, K. M. (2008). Perpetration of violence, violent victimization, and severe mental illness: Balancing public health concerns. *Psychiatric Services, 59*(2), 153-164. doi: 10.1176/appi.ps.59.2.153

Coid, J., Hickey, N., Kahtan, N., Zhang, T., & Yang, M. (2007). Patients discharged from medium secure forensic psychiatry services: Reconvictions and risk factors. *The British Journal of Psychiatry, 190*, 223-229. doi: 10.1192/bjp.bp.105.018788

Cooke, D. J., & Michie, C. (1999). Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology, 108*(1), 58–68. doi: 10.1037/0021-843X.108.1.58

Cullen, F. T., Gendreau, P., Jarjoura, C. R., & Wright, J. P. (1997). Crime and the Bell Curve: Lessons from intelligent criminology. *Crime and Delinquency, 43*(4), 387-411. doi: 10.1177/0011128797043004001

Davies, S., Clarke, M., Hollin, C., & Duggan, C. (2007). Long-term outcomes after discharge from medium secure care: A cause for concern. *The British Journal of Psychiatry, 191*, 70-74. doi: 10.1192/bjp.bp.106.029215

Dienst voor het Strafrechtelijk beleid. (2015). *Veroordelingen, opschortingen en interneringen*. Retrieved from http://www.dsb-spc.be/web/index.php?option=com_content&task=view&id=28&Itemid=47&lang=dutch

Dirks-Linhorst, P. A., & Linhorst, D.M. (2010). Recidivism Outcomes for Suburban Mental Health Court Defendants. *American Journal of Criminal Justice, 37*(1), 76-91. doi: 10.1007/s12103-010-9092-0

Dirks-Linhorst, P.A, & Linhorst, D.M. (2012). Recidivism outcomes for suburban mental health court defendants. *American Journal of Criminal Justice, 37*(1), 76-91. doi: 10.1007/s12103-010-9092-0

Dolan, M., & Khawaja, A. (2004). The HCR–20 and post-discharge outcome in male patients discharged from medium security in the UK. *Aggressive Behavior, 30*(6), 469-483. doi: 10.1002/ab.20044

Ducro, C, & Pham, T. (2006). Evaluation of the SORAG and the Static-99 on Belgian sex offenders committed to a forensic facility. *Sexual Abuse: A Journal of Research and Treatment, 18*(1), 15-26. doi: 10.1007/s11194-006-9003-6

Edwards, J., Steed, P., & Murray, K. (2002). Clinical and forensic outcome 2 years and 5 years after admission to a medium security unit. *The Journal of Forensic Psychiatry, 13*(1), 68-87.

Falshaw, L., Bates, A., Patel, V., Corbett, C., & Friendship, C. (2003). Assessing reconviction, reoffending and recidivism in a sample of UK sexual offenders. *Legal and Criminological Psychology, 8*(2), 135-268. doi: 10.1348/135532503322362979

Friendship, C., McClintock, T., Rutter, S., & Maden, A. (1999). Re-offending: Patients discharged from a regional secure unit. *Criminal Behaviour and Mental Health, 9*(3), 226-236. doi: 10.1002/cbm.315

Gilleir, F., Easton, M., & Ponsaers, P. (2010). Een U-Turn voor jongvolwassen veelplegers. *Panopticon, 31*(1), 5-19.

Gray, N. S., Taylor, J., & Snowden, R. J. (2008). Predicting violent reconvictions using the HCR-20. *The British Journal of Psychiatry, 192*(5), 384-387. doi: 10.1192/bjp.bp.107.044065

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist*. Toronto, ON: Multi-Health Systems.

Hurducas, C. C., Singh, J. P., Ruiter, C. de, & Petrila, J. (2014). Violence risk assessment tools: A systematic review of surveys. *International Journal of Forensic Mental Health, 13*(3), 181-192. doi: 10.1080/14999013.2014.942923

Jeandarme, I., Wittouck, C., Laenen, F. Vander, Varé, J. De, Grouwels, Y., Ampe, M., & Oei, T. I. (2013). *Disciplinary infractions in three Flemish forensic institutions: Prevalence and judicial reaction.* Paper presented at the 8th European Congress on Violence in Clinical Psychiatry, Ghent, Belgium.

Jeandarme, I., Wittouck, C., Laenen, F. Vander, Ampe, M., Grouwels, Y., Varé, J. De, . . . Bogaerts, S. (2015). Forensische pilootprojecten 'medium security': Incidenten tijdens behandeling van geïnterneerden ressorterend onder CBM Gent [Forensic 'medium security' projects: registration of incidents during treatment of forensic psychiatric patients of the Commission of Social Defense Ghent]. *Panopticon, 36*(1), 26-45.

Joyal, C., Dubreucq, J., Gendron, C., & Millaud, F. (2007). Major mental disorders and violence: A critical update. *Current Psychiatry Reviews, 3*(1), 33-50. doi: 10.2174/157340007779815628

Kennedy, H. G. (2002). Therapeutic uses of security: Mapping forensic mental health services by stratifying risk. *Advances in Psychiatric Treatment, 8*, 433-443. doi: 10.1192/apt.8.6.433

Khiroya, R., Weaver, T., & Maden, T. (2009). Use and perceived utility of structured violence risk assessments in English medium secure forensic units. *Psychiatric Bulletin, 33*, 129-132. doi: 10.1192/pb.bp.108.019810

Kindness, A., Kim, H., Alder, S., Edwards, A., Parekh, A., & Olson, L. M. (2009). Court compliance as a predictor of postadjudication recidivism for domestic violence offenders. *Journal of Interpersonal Violence, 24*(7), 1222-1238. doi: 10.1177/0886260508322197

Laenen, F. Vander, & Cauwer, B. De (2011). Just-care, justitie en zorg in dialoog over internering. *Panopticon, 32*(3), 55-59.

Lee, D. T. (2003). Community-treated and discharged forensic patients: an 11-year follow-up. *International Journal of Law and Psychiatry, 26*(3), 289-300. doi: 10.1016/s0160-2527(03)00039-6

Leistico, A. M., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the hare measures of psychopathy to antisocial conduct. *Law and Human Behavior, 32*(1), 28-45. doi: 10.1007/s10979-007-9096-6

Lievens, J. (2001). Werken met secundaire data. In J. Billiet & H. Waege (Eds.), *Een samenleving onderzocht: Methoden van sociaal-wetenschappelijk onderzoek* (pp. 343-363). Antwerpen: Standaard Uitgeverij.

Luettgen, J., Chrapko, W. E., & Reddon, J. E. (1998). Preventing violent re-offending in Not Criminally Responsible patients: An evaluation of a continuity of treatment program. *International Journal of Law and Psychiatry, 21*(1), 89-98. doi: 10.1016/S0160-2527(97)00051-4

Lund, C., Hofvander, B., Forsman, A., Anckarsater, H., & Nilsson, T. (2013). Violent criminal recidivism in mentally disordered offenders: A follow-up study of 13-20 years through different sanctions. *International Journal of Law and Psychiatry, 36*(3-4), 250-257. doi: 10.1016/j.ijlp.2013.04.015

Lunsky, Y., Gracey, C., Koegl, C., Bradley, E., Durbin, J., & Raina, P. (2011). The clinical profile and service needs of psychiatric inpatients with intellectual disabilities and forensic involvement. *Psychology, Crime & Law, 17*(1), 9-23. doi: 10.1080/10683160903392277

Maden, A., Scott, F., Burnett, R., Lewis, G. H., & Skapinakis, P. (2004). Offending in psychiatric patients after discharge from medium secure units: Prospective national cohort study. *British Medical Journal, 328*(7455), 1534. doi: 10.1136/bmj.38075.467569.EE

Maes, B., Goethals, J., & Verlinden, S. (2009). *Personen met een verstandelijke handicap onderhevig aan een interneringsmaatregel*. Leuven: Steunpunt Welzijn, Volksgezondheid en Gezin.

Martin, M. S., Dorken, S. K., Wamboldt, A. D., & Wootten, S. E. (2012). Stopping the revolving door: A meta-analysis on the effectiveness of interventions for criminally involved individuals with major mental disorders. *Law and Human Behavior, 36*(1), 1-12. doi: 10.1037/h0093963

Menghini, M., Ducro, C., & Pham, T. H. (2005). Evaluation de la récidive auprès d'agresseurs sexuels issus d'un hôpital sécuritaire en Belgique francophone [Evaluation of recidivism among sex offenders from a secure hospital in the French part of Belgium]. *Psychiatrie & Violence, 51*(1), 39-48.

Moens, I., & Pauwelyn, L. (2012). *Geen opsluiting, maar sleutels tot re-integratie. Voorstellen voor een gecoördineerd zorgtraject voor geïnterneerden*. Brussel: Zorgnet Vlaanderen.

Monahan, J., Steadman, H., Silver, E., Appelbaum, P., Robbins, P., Mulvey, E., . . . Banks, S. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York, NY: Oxford University Press.

Morgan, R. D., Flora, D. B., Kroner, D. G., Mills, J. F., Varghese, F., & Steffan, J. S. (2012). Treating offenders with mental illness: A research synthesis. *Law and Human Behavior, 36*(1), 37-50. doi: 10.1037/h0093964

Müller-Isberner, R., Freese, R., Jockel, D., & Gonzalez Cabeza, S. (2000). Forensic psychiatric assessment and treatment in Germany. Legal framework, recent developments, and current practice. *International Journal of Law and Psychiatry, 23*(5-6), 467-480.

Murrie, D. C., Boccaccini, M. T., Caperton, J., & Rufino, K. (2012). Field validity of the Psychopathy Checklist-Revised in sex offender risk assessment. *Psychological Assessment, 24*(2), 524-529. doi: 10.1037/a0026015

Nilsson, T., Wallinius, M., Gustavson, C., Anckarsäter, H., & Kerekes, N. (2011). Violent recidivism: A long-time follow-up study of mentally disordered offenders. *PLoS ONE, 6*(10), e25768. doi: 10.1371/journal.pone.0025768

Nordstrom, A., & Kullgren, G. (2003). Victim relations and victim gender in violent crimes committed by offenders with schizophrenia. *Social Psychiatry and Psychiatric Epidemiology, 38*(6), 326-330. doi: 10.1007/s00127-003-0640-5

Nowak, M. J., & Nugter, M. A. (2014). 'Ik zit mijn tijd wel uit...' II. Een replicatieonderzoek naar recidiven bij patiënten met een strafrechtelijke machtiging ['I'll serve my time...' II. A replication study of recidivism in patients with a one-year hospital order]. *Tijdschrift voor Psychiatrie, 56*(4), 237-246.

Pham, T. H, & Ducro, C. (2008). Évaluation du risque de récidive en Belgique francophone: Données préliminaires d'analyse factorielle de la "Sex Offender Recidivism Appraisal Guide" (SORAG) et de la Statique-99 [Assessing recidivism risk in the French-speaking Belgium: Preliminary data factor analysis of the "Sex Offender Recidivism Appraisal Guide" (SORAG) and Static-99]. *Annales Médico-Psychologiques, 166*(7), 575-579. doi: 10.1016/j.amp.2008.06.001

Pham, T. H, Ducro, C, Marghem, B, & Réveillère, C. (2005). Evaluation du risque de récidive au sein d'une population de délinquants incarcérés ou internés en Belgique francophone [Assessing recidivism risk in a population of incarcerated or interned in French-speaking Belgium]. *Annales Médico-Psychologiques, 163*(10), 878-881. doi: 10.1016/j.amp.2005.09.013

Philipse, M., Ruiter, C. de, Hildebrand, M., & Bouman, Y. (2000). *HCR-20. Beoordelen van het risico van gewelddadig gedrag* (Versie 2). Nijmegen/Utrecht: Pompestichting/Dr. Henri van der Hoeven Stichting.

Rice, M. E., Harris, G. T., & Lang, C. (2013). Validation of and revision to the VRAG and SORAG: The Violence Risk Appraisal Guide-Revised (VRAG-R). *Psychological Assessment, 25*(3), 951-965. doi: 10.1037/a0032878

Robert, L. (2014). Wat werkt in Nederland en wat niet? Een meta-analyse van Nederlands recidiveonderzoek naar de effecten van strafrechtelijke interventies. *Panopticon, 35*(5), 480-485.

Robert, L., & Maes, E. (2012). *Wederopsluiting na vrijlating uit de gevangenis*. Brussel: NICC.

Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J., Reuter, P., & Shawn, D. (1998). *Preventing crime: What works, what doesn't, what's promising*. Washington, DC: National Institute of Justice, US Department of Justice.

Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., . . . Otto, R. K. (2014). International perspectives on the practical application of violence risk assessment: A global survey of 44 countries. *International Journal of Forensic Mental Health, 13*(3), 193-206. doi: 10.1080/14999013.2014.922141

Skipworth, J., Brinded, P., Chaplow, D., & Frampton, C. (2006). Insanity acquittee outcomes in New Zealand. *Australian and New Zealand Journal of Psychiatry, 40*(11-12), 1003-1009. doi: 10.1111/j.1440-1614.2006.01924.x

Smedt, A. De, Mariën, L., & Vermeiren, R. (2008, Semptember). *Profilering van de medium risk geïnterneerden in forensische psychiatrische settings.* Paper presented at the Fourth Vlaams Geestelijk Gezondheidscongres, Ghent, Belgium.

Spreen, M., Brand, E., Horst, P. Ter, & Bogaerts, S. (2014). *Handleiding en methodologische verantwoording HKT-R, Historisch, Klinische en Toekomstige-Revisie*. Groningen: Dr. S. van Mesdag.

Steadman, H. J., Mulvey, E. P., Monahan, J., Robbins, P., Appelbaum, P., Grisso, T., … Silver, E. (1998). Violence by people discharged from acute psychiatric inpatient facilities and by others in the same neighborhoods. *Archives of General Psychiatry, 55*(5), 393-401. doi: 10.1001/archpsyc.55.5.393

Tabita, B., Santi, M. G. de, & Kjellin, L. (2012). Criminal recidivism and mortality among patients discharged from a forensic medium secure hospital. *Nordic Journal of Psychiatry, 66*(4), 283-289. doi: 10.3109/08039488.2011.644578

To, W. T., Vandevelde, S., Soyez, V., Smet, S. De, Boers, A., & Vanheule, S. (2014). Treatment perspectives on interned mentally ill offenders in a forensic psychiatric center (FPC): A Delphi study on experts' opinions. *Psychology, Crime and Law, 20* (1), 61-77. doi: 10.1080/1068316X.2012.736511

Triest, K. (2009). *Internering en gedwongen opname*: Masterproef, Faculteit Rechten, Universiteit Gent.

Vertommen, H., Verheul, R., Ruiter, C. de, & Hildebrand, M. (2002). *Handleiding bij de herziene versie van Hare's Psychpathie Checklist*. Lisse: Swets Test Publishers.

Vitacco, M. J., Rybroek, G. J. Van, Erickson, S. K., Rogstad, J. E., Tripp, A., Harris, L., & Rodney, M. (2008). Developing services for insanity acquittees conditionally released into the commmunity: Maximizing success and minimizing recidivism. *Psychological Services, 5*(2), 118-125. doi: 10.1037/1541-1559.5.2.118

Vojt, G., Thomson, L. D. G., & Marshall, L. A. (2013). The predictive validity of the HCR-20 following clinical implementation: Does it work in practice? *The Journal of Forensic Psychiatry & Psychology, 24*(3), 371-385. doi: 10.1080/14789949.2013.800894

Wartna, B. S. J. (2009). *In de oude fout. Over het meten van recidive en het vaststellen van het succes van strafrechtelijke interventies*. Den Haag: WODC.

Wartna, B. S. J., Harbachi, S. el, & Knaap, L. M. van der (2005). *Buiten behandeling: Een cijfermatig overzicht van de strafrechtelijke recidive van ex-terbeschikkinggestelden*. Den Haag: WODC/Boom Juridische uitgevers.

Wartna, B. S. J., & Nijssen, L. T. J. (2006). *National studies on recidivism: An inventory of large-scale recidivism research in 33 European countries*. Den Haag: WODC.

Wartna, B. S. J., & Tollenaar, N. (2004). *Bekenden van Justitie. Een verkennend onderzoek naar de 'veelplegers' in de populatie van vervolgde daders*. Den Haag: WODC.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence* (Version 2). Burnaby, BC: Simon Fraser University.

Wechsler, D. (2000). *WAIS-III Nederlandstalige bewerking. Afname en scoringshandleiding*. Lisse: Swets.

Yiend, J., Chambers, J. C., Burns, T., Doll, H., Fazel, S., Kaur, A., . . . Fitzpatrick, R. (2011). Outcome measurement in forensic mental health research: An evaluation. *Psychology, Crime & Law, 17*(3), 277-292. doi: 10.1080/10683160903203938

Yoshikawa, K., Taylor, P. J., Yamagami, A., Okada, T., Ando, K., Taruya, T., . . . Kikuchi, A. (2007). Violent recidivism among mentally disordered offenders in Japan. *Criminal Behaviour and Mental Health, 17*(3), 137-151. doi: 10.1002/cbm.652

Who are the victims of NGRI acquittees? A study on Belgian internees

Submitted

12

Jeandarme, I., Vandenbosch, L., Groenhuijsen, M., Oei, T. I., & Bogaerts, S.

ABSTRACT

Background. The limited literature on victim characteristics of offenders found not guilty by reason of insanity (NGRI) shows that most victims are adults who are known to the offender. It is currently unclear whether victims are mainly male or female or whether there are differences in the type of victims according to the offenders' psychiatric disorder.

Method. Victim characteristics were retrospectively collected from 362 NGRI acquittees, and the influence of psychiatric diagnosis on victim profiles was examined.

Results. Victims were mainly adult acquaintances and were equally likely to be male or female. Family members and caregivers were the most frequent type of acquaintance victims. Further analyses suggested that these victim characteristics were similar for perpetrators with different psychiatric diagnoses.

Conclusion. Victimization of strangers and minors was unlikely in NGRI offenders.

INTRODUCTION

Studying victim characteristics in mentally disordered offenders (MDOs) enables effective treatment and the prevention of future victims (Mezey, 2007). In contrast to the scholarly attention to MDOs, however (e.g., Bowers et al., 2011; Coid, Kahtan, Gault, Cook, & Jarman, 2001), studies have rarely examined the characteristics of their victims. Moreover, no study thus far has addressed the characteristics of victims of MDOs found not guilty by reason of insanity (NGRI) after being accused of a crime. This is surprising because studies have shown that there are more victims of NGRI acquittees compared to criminally responsible offenders (e.g.,1.5 vs. 0.7 in Gulayets, 2016). Moreover, violence committed by people with severe mental disorders cannot simply be explained by their psychiatric symptoms. Violence is rather embedded within the social circumstances, life experiences, and relationships of the perpetrator.

Victim gender and age

The mean age of victims in NGRI offender populations has rarely been reported. One study found that the vast majority (95.7%) of NGRI acquittee cases involved adult victims (Gulayets, 2016). Menezes, Oyebode, and Haque (2009) found a similar trend in Zimbabwe, although they did report a larger minority of minor victims (i.e., 24% of the victims were under 16 years old). These findings are in line with studies focusing on more broadly defined MDO populations, which found prevalence rates of minor offenders to be below 20% on average (e.g., Gradillas, Williams, Walsch, & Fahy, 2007; Nordstrom & Kullgren, 2003b). However, in NGRI sex offenders, 50% of the victims were minor victims (Novak, McDermott, Scott, & Guillory, 2007).

Research examining the gender of victims of NGRI offenders has shown mixed results. Crocker et al. (2015) examined victim gender in a Canadian[1] sample and found that victims were males in slightly more than half of the index offenses involving crimes against a person (53.3%), whereas Gulayets (2016) found the opposite (55.2% female victims). Menezes, Oyebode, and Haque (2009) also found that the total number of female victims (55.4%) was greater than the total number of male victims (44.6%). According to Cirincione, Steadman, and McGreevy (1995), insanity acquittals were more likely in cases of male victims than female victims.

Offender-victim relationship

Most research has focused on the relationship between victims and NGRIs. In terms of the victims' relationship with the accused, 26.7% of the total number of victims were strangers, 21.1% were acquaintances, 16.4% were family members, and 13.1% were professionals (for example, police officers and mental health staff) (Livingston, Wilson, Tien, & Bond, 2003). In contrast, Crocker et al. (2015) reported that family members (including partners) (33.7%) in a Canadian sample were the most targeted victims of index offenses against a person, followed by professionals (22.9%), strangers (22.7%),and other people known to the accused (20.7%). Crocker, Seto, Nicholls, and Côté (2013) focused on very serious offenses only and found differences in the distribution of accused people's relationships with the victims according to the type of index offense. Particularly in cases of murder or attempted murder, the victims were most often family members (57%), with parents being the most likely victims followed by partners or spouses. Victims of sexual offenders were more likely to be strangers (43%) and less likely to be a family member (16.7%). Studies comparing NGRI acquittees to offenders found criminally responsible confirmed that the chance of NGRI acquittal was associated with acquaintance victims (Cirincione et al., 1995; Parker, 1987). However, one study that found a larger number of stranger victims (42.9%) failed to find a significant difference between the two groups (Gulayets, 2016).

Taken together, it seems that studies on NGRI populations generally agree that there are more acquaintance victims than stranger victims, which is in line with research on more broadly defined MDO populations (e.g., Gow, Choo, Darjee, Gould, & Steele, 2010; Gradillas et al., 2007). As noted by Swanson et al. (1998), social contact is a mixed blessing for people with a severe mental illness. Particularly for those with extreme psychiatric impairment, frequent contact with others may produce conflict, stress, and opportunities for physical violence. Individuals in a close relationship with the psychiatric patient are more likely to be assaulted. As such, several studies have reported a quite substantial proportion of relatives (i.e., family members and intimate partners) among the victims of NGRI (Crocker et al., 2013; Crocker et al., 2015; Menezes et al., 2009; Parker, 1987). Moreover, researchers who identified a so-called key cluster of risk factors of becoming a victim have given a central focus on relatives. This cluster includes being a mother or immediate co-habiting relative of an individual who is financially dependent and diagnosed with schizophrenia or impaired by substance abuse and a low user or non-user of mental health

services (Estroff, Swanson, Lachicotte, Swartz, & Bolduc, 1998). Even in cases of repeated violence before and after hospitalization, the target was the same person in the majority of the cases, most often a spouse, an intimate relative, or other family member (Tardiff, Marzuk, Leon, & Portera, 1997).

For obvious reasons, different victim profiles can be expected during hospitalization since inpatients have fewer opportunities to assault relatives while institutionalized. Although the NGRI offender literature has not explored hospitalization in more detail, results seem to suggest that at least in some cases, the victims involved are mental health professionals or co-patients (e.g., Crocker et al., 2015). Most inpatient studies on more broadly defined MDOs found that hospital staff in particular were victimized, as were other patients but to a lesser extent (e.g., Gow et al., 2010; Weizmann-henelius & Suutala, 2000). In addition, hospital staff, nurses, and ward staff members (i.e., individuals who have the closest contact with the patient) were more frequently victimized than clinical or supervisory staff (e.g., Kelly, Subica, Fulginiti, Brekke, & Novaco, 2015; Nicholls, Brink, Greaves, Lussier, & Verdun-Jones, 2009).

However, not all studies have consistent findings regarding which type of acquaintance victims (hospital staff or other patients) are the most likely to become victims of inpatient MDOs. For example, some studies found that it was equally likely (Nicholls et al., 2009) or even more likely (Bader, Evans, & Welsh, 2014; Daffern, Mayer, & Martin, 2003) that victims were co-patients than staff. In addition, violent incidents toward staff are more likely to be formally reported than those directed to other patients. In line with these findings, scholars generally agree that patients with a severe mental illness run a higher risk of victimization than the regular population (Kamperman et al., 2014; Maniglio, 2009). In addition, inpatients are regarded as having higher risk of victimization than outpatients (de Mooij et al., 2015).

Type of psychiatric disorder

To the best of our knowledge, there is no research on NGRI offenders investigating the victim-offender relationship in regard to the psychiatric disorder of the offender. In MDOs, Johnston and Taylor (2003) found that regardless of the type of personality disorder, the diagnosis of a personality disorder was positively associated with stranger victimization. This finding was not confirmed in another study (Goethals, Gaertner, Buitelaar, & van Marle, 2008). Goethals et al.

(2008) investigated whether victims of psychotic patients (with or without a combined personality disorder) detained under the Dutch Entrustment Act would be more likely to be a relative, friend, or acquaintance than the victims of patients with only a personality disorder, but they found no significant difference. So far, no study has explored whether the type of psychiatric disorder relates to differences in the gender or age of victims.

In sum, the literature suggests that victims of NGRI offenders are in many ways similar to the victims of MDO offenders; i.e., they are likely to be adult victims who are known to the offender and they are equally likely to be male or female. However, the available research is still limited in its scope and has rarely included in-depth analyses.

The present study

This study investigates information on victims of interpersonal violence committed by NGRI offenders. The study uses both categorical and continuous data, and it relates victim characteristics to offender characteristics. First, information is given regarding the relationship to the offender and the circumstances of the index offense. Second, based on prior literature, we test the following hypotheses:

- *Hypothesis 1*. More adult victims than minor victims occur.
- *Hypothesis 2*. More acquaintance victims than stranger victims occur.

We also examine whether differences exist according to the nature of the psychiatric disorder of the offender, the gender and age of the victims, and their relationship with the offender. As such, we pose the following research questions:

- *Research Question 1*. Is there a difference between the number of female and male victims?
- *Research Question 2*. Is there a difference regarding gender, age, and relationship of the victims after stratification for the psychiatric diagnoses of the NGRIs (psychotic vs. personality disorder vs. psychotic and personality disorder vs. other)?

METHOD

Setting and participants

This multicenter study was conducted at three medium security units located in the communities of Bierbeek, Zelzate, and Rekem in Flanders (the Dutch-speaking part of Belgium). Medium security units provide a treatment setting for NGRI offenders who do not require care in a high secure hospital but are considered unsuitable for a general psychiatric ward or outpatient care. In Belgium, several conditions have to be fulfilled before one can be acquitted because of a mental disorder. First, the offender must have committed a felony or misdemeanor for which the criminal law sets a minimum penalty of at least eight days. Second, the person must be found unfit to control his actions due to his mental disturbance and must be found socially dangerous at the moment of sentencing. The Belgian system is dichotomized between criminally responsible and irresponsible people without further differentiation between complete irresponsibility and severely diminished responsibility. Upon NGRI acquittal, socially dangerous people are submitted to an internment measure under the supervision of a regional court, the Commission for the Protection of Society (CPS). The aim of this measure is to protect society form further offenses and provide the necessary treatment for the offender (see Jeandarme, Habets, Oei, and Bogaerts (2016) for a description of the Belgian forensic psychiatric system).

All participants (*N* = 362) were found NGRI after having committed a violent offense. Participants were mostly male patients (94.5%, *n*= 342). Most of the patients had the Belgian nationality (three missing; 89.4%, *n* = 321). The mean age at first admission to a medium security unit was 36.3 years (*SD* = 11.04, range = 18.8–73.4 years). The most common diagnoses were personality disorders (74.6%, *n* = 270), substance use disorders (55.8%, *n* = 202) and psychotic disorders (40.3%, *n* = 146) according to the Diagnostic and Statistical Manual of mental disorders-IV-Text Revision (DSM-IV-TR; American Psychiatric Association, 2000).

Measures and procedures

A retrospective analysis was done to examine a consecutive cohort admitted to one of the medium security units in Flanders during the period of 2001–2010. Eleven NGRI offenders refused to give passive consent, and 159 admitted offenders had an index offense without

victims, such as theft. The other 373 NGRI offenders with violent index offenses against people were included. The study population comprised 68.6% of the total admitted population during that period. Data from eleven patients were not analyzed because there was no information available on the victim characteristics, leaving a total study population of *N* = 362.

Information on the violent index offenses was retrieved from the Central Criminal Records of the Ministry of Justice and cross referenced with the different court administrations. Violent offenses were restricted to violence toward another person, referring to the intentional use of physical force or power – threatened, attempted, or actual – against another person. For each NGRI offender, victim characteristics regarding demographic characteristics and the relationship with the offender were analyzed. Information on NGRI offenders was gathered through hospital files regarding demographics, type of offense, status of hospitalization, and psychiatric DSM diagnoses. Ethical approval was obtained from the Medical Ethical Commission of the University Hospital of Antwerp.

Demographic characteristics of victims. The gender of the victim (male or female) and whether the victim was a minor or adult were recorded. The total number of victims, number of female and male victims, and the number of adult and minor victims were also noted.

Victim-offender relationship. First, information was recorded about whether the victim was a stranger or an acquaintance of the offender. *Stranger victims* were defined as victims not known to the offender 24 hours prior to the offense, while all other victims were defined as *acquaintance victims*. The acquaintance victims were divided into three groups: care victims, family victims, and other victims. *Care victims* consisted of 1) caregivers, who are broadly defined as anyone involved in the treatment or confinement of the NGRI offenders, such as nurses, doctors, or prison officers; and 2) co-residents, which refer to other psychiatric patients or other inmates. *Family victims* consisted of 1) intimate partners defined as current or former romantic partners and 2) family victims referring to close family members such as parents and siblings. *Other victims* comprised 1) work-related victims referring to colleagues working with the NGRI offenders at the time of the offense, 2) friends, and 3) all other victims, such as instance neighbors or children of the intimate partner. The total number of victims in each category was calculated.

Type of committed index offense. The type of violence committed against each victim was divided into homicide/attempted homicide, sexual hands-on assault, and other violent offenses.

Demographic information of the offender. Age, gender, and nationality were coded.

Hospitalization status of the offender. We examined whether the patient was *inpatient* in an institution (hospital, prison) or *outpatient* while committing the index offense.

Psychiatric DSM diagnoses of the offender. Psychiatric DSM diagnoses were clustered into four categories: 1) psychotic disorders without personality disorder (19.6%, *n* = 71), 2) personality disorders without psychotic disorder (53.9%, *n* = 195), 3) psychotic disorders in combination with personality disorders (20.7%, *n* = 75), and 4) patients with diagnoses other than personality or psychotic disorders (5.8%, *n* = 21).

Analytical strategy

SPSS Version 22 was used for the descriptive statistical analyses. Repeated measures ANOVA was used for parametric data after removal of outliers in key variables based on their value of two times the standard deviation above the mean and root square transformation (Miller, 1991; Osborne, 2010). Post hoc comparisons were done using Bonferroni correction. Outliers were omitted for only the age variable. To ensure that omitting these participants would not bias the analyses, non-parametric exploratory analyses were also conducted with all participants. These analyses demonstrated similar results and are thus not reported.

Repeated measures ANOVA is considered a rather robust technique, and large samples suffer less from unequal variances of the recommended skewness and kurtosis levels (Tabachnick & Fidell, 2013). However, literature suggests that the normality distribution could be improved (Osborne, 2010). The data were considered to approach a normality distribution when the absolute skewness level was ≤3 and the kurtosis level was ≤10 (Kline, 2010; Tabachnick & Fidell, 2013).

RESULTS

Descriptive statistics

On average, 2.2 victims (*SD* = 2.46, range 1–23) were registered, which led to a total of 792 identified unique victims. There was missing information regarding the victim's age (2.6%, *n* = 20), gender (7.5%, *n* = 58), and relationship to the offender (6.2%, *n* = 48). Regarding the victim's age,

most NGRI offenders (87.8%) exclusively victimized adult victims, whereas a few (7.5%) had only minor victims. A minority victimized both adult and minor victims (4.7%). On average, 1.8 (*SD* = 1.93, range = 0–18) adult victims and 0.3 minor victims (*SD* = 1.70, range = 0–23) were identified, which led to a total of 663 adult and 109 minor victims.

Regarding the victims' gender, most NGRIs (42.9%) exclusively victimized female victims, whereas 34.6% had only male victims, and 22.6% had both female and male victims. On average, 0.9 (*SD* = 1.50, range = 0–13) male victims and 1.1 (*SD* = 1.70, range = 0–20) female victims were found for the NGRIs, which led to a total of 342 male and 392 female victims.

Regarding the victim's relationship, most NGRIs (66.4%) exclusively victimized acquaintance victims, whereas a minority (26.6%) had only stranger victims, and 7.0% had both stranger and acquaintance victims. The analyses showed an average of 1.4 (*SD* = 1.90, range = 0–20) acquaintance victims and 0.6 stranger victims (*SD* = 1.68, range = 0–23), which led to a total of 513 identified acquaintance victims (21.3% family victims) and 231 stranger victims. The distribution of the victim profile according to the nature of the perpetrated violence is shown in Table 1. A minority of the offenders threatened or assaulted the victims while they were inpatients[i] (9.4%, *n* = 34), but the majority were outpatients (90.6%, *n* = 328).

Differences according to victim age

Total group. To test the first hypothesis, a repeated measures ANOVA was conducted with the parametric data (*n* = 358). The result showed a significant difference between the number of adult victims (*M* = 1.22, *SD* = 0.59) and minor victims (*M* = 0.13, *SD* = 0.37), $F(1, 357) = 655.27$, $p < .001$, $\eta p = .65$.

Stratification according to diagnoses. To answer the second research question, we used repeated measures ANOVA (*n* = 358) to test whether this difference would be moderated by patients' psychiatric diagnoses. The results showed this was not true ($F(3, 354) = 1.94$, $p = .12$, partial Eta squared = .01). No influence of diagnoses was thus found, as more adult than minor victims occurred in all groups.

Table 1. Distribution of victim gender, age and relationship for the total number of offenses and stratified according to the type of violent offense

	Index offense total (*n* = 792)		Homicide (*n* = 105)		Sexual assault (*n* = 119)		Other violence (*n* = 568)	
	N	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Gender								
Male	342	46.6	48	46.2	41	35.3	253	49.2
Female	392	53.4	56	53.8	75	64.7	261	50.8
Unknown	58		1		3		54	
Age								
Adult	663	85.9	93	88.6	30	25.2	540	98.5
Minor	109	14.1	12	11.4	89	74.8	8	1.5
Unknown	20		0		0		20	
Relationship to offender								
Unknown	48		2		0		46	
Stranger victim	231	31.0	16	15.5	43	36.1	172	33.0
Acquaintance victim	513	69.0	87	84.5	76	64.7	350	67.0
Care victims	167/513	32.6	10/87	11.5	21/76	27.6	136/350	38.9
Patients/convicts	30/513	5.8	6/87	6.9	1/76	1.3	23/350	6.6
Care givers/peace officers	137/513	26.7	4/87	4.6	20/76	26.3	113/350	32.3
Family members	191/513	37.2	52/87	59.8	25/76	32.9	114/350	32.6
Blood relatives	86/513	16.8	22/87	25.3	15/76	19.7	49/350	14.0
Intimate partners	105/513	20.5	30/87	34.5	10/76	13.2	65/350	18.6
Other acquaintance victims	155/513	30.2	25/87	28.7	30/76	39.5	100/350	28.6
Friends	28/513	5.5	8/87	9.2	6/76	7.9	14/350	4.0
Work-related	10/513	1.9	0	0	0	0	10/350	2.9
Rest category	117/513	22.8	17/87	19.5	24/76	31.6	76/350	21.7

Differences according to victims' gender

Total group. Repeated measures ANOVA was conducted to test the first research question (n = 362). The result showed no significant difference between the number of male victims (M = 0.68, SD = 0.70) and female victims (M = 0.78, SD = 0.69, $F(1, 361) = 2.85$, p = .09).

Stratification according to diagnoses. To answer the second research question, a repeated measures ANOVA (n = 362) tested whether the difference in male versus female victims would be

moderated by patients' psychiatric diagnoses. The results could not confirm this assumption (F (3, 358) = 1.71, p = .19). No effect of diagnosis was thus found.

Differences according to victims' relationship with the perpetrator

Total group. Repeated measures ANOVA was conducted (n = 362) showed a significant difference between the number of acquaintance victims (M = 0.95 SD = 0.72) and stranger victims (M = 0.42, SD = 0.68, F(1, 361) = 63.57, $p < .001$, $\eta p^2 = .15$).

Stratification according to diagnoses. To answer the second research question, repeated measures ANOVA (n = 362) tested whether this difference would be moderated by patients' psychiatric diagnoses. The results could not confirm this assumption (F(3, 358) = 1.65, p = .18). No effect of diagnosis was thus found.

Differences according to most occurring acquaintance victim

Total group. To test the second research question, repeated measures ANOVA tested which would be the most frequent group of acquaintance victims. The test results showed that there was a significant difference between family victims (M = 0.59, SD = 0.62), care victims (M = 0.43, SD = 0.67), and other victims (M = 0.40, SD = 0.66), F(2, 522) = 4.99, $p < .001$, $\eta p^2 = .02$. Post hoc tests with Bonferroni correction were used to interpret this finding and revealed that there was a significant difference between the number of family victims compared to care and other victims, while there was no significant effect between the care and other victims.

Stratification according to diagnoses. To answer the second research question, repeated measures ANOVA (n = 362) tested whether this difference would be moderated by patients' psychiatric diagnoses. The results could not confirm this assumption (F(6, 516) = 1.28, p = .27). No effect of diagnosis was thus found.

DISCUSSION

This study examined the characteristics (age, gender, and relationship to the offender) of victims of violence in cases of NGRI and investigated whether the psychiatric diagnosis of NGRI offenders would influence the findings. The results were in line with previous research describing that adult

victims known to the offender were most frequently victimized. Our study thus further adds to the evidence that stranger victimization is unlikely in cases of NGRI. Furthermore, diagnostic categories did not moderate these findings.

Cases of NGRI in this study involved more adult victims than minor victims. This finding is in line with the observations of previous research (e.g., Gulayets, 2016) and confirms the first hypothesis. This increased number of adult victims was found for cases of NGRI in each diagnostic category. In contrast, when specifically investigating sexual assault, three-fourths of the victims were minors. This finding is in line with another NGRI study that focused on sex offenders. Together, the literature highlights the need to differentiate according to the nature of the perpetrated violence (Novak et al., 2007).

As for gender, no clear pattern emerged in the present study, which is consistent with previous mixed findings on gender (Crocker et al., 2015; Gulayets, 2016). No gender differences were found for the total group, and there was no interaction effect with diagnosis. When specifically looking at sexual assault, less than one-third of the victims were male, which may suggest a general tendency of men to underreport such crimes. Men in general have a lower likelihood of reporting victimization because of stigma, shame, or fear of having their sexuality questioned (Davies, 2002; Tewksbury, 2007). Taken together, results from prior research and the current study seem to suggest that gender is not clearly related to victimization. This observation is also in line with more broadly defined MDO populations (e.g., Liettu, Saavala, Hakko, Rasanen, & Joukamaa, 2009; Nordstrom & Kullgren, 2003b). However, in the general population, victims of violence tend to be more often men (Kamperman et al., 2014). Further research is needed to study separate forms of victimization such as sexual assault more in depth.

The current study further adds to the evidence that victims of violence in cases of NGRI are more often acquaintances than strangers (Crocker et al., 2015; Livingston et al., 2003). This is also consistent with results from more broadly defined MDO populations (Estroff et al., 1998; Gow et al., 2010; Gradillas et al., 2007) and confirms our third hypothesis. Being in a close relationship with an MDO increases the likelihood of victimization. This became evident when investigating the type of victim acquaintance: family members were more commonly victimized than the other groups. More specifically, in the case of manslaughter/homicide, 59.8% of the acquaintance victims were family members, whereas 32.6% of the acquaintance victims were family members

in the less serious violent offenses. This finding is consistent with other research, which found that family victims were injured more severely than other victims (Crocker et al., 2013). According to Nordstrom and Kullgren (2003a), this finding can be explained by the higher threshold for reporting crimes within the family and stronger emotional bonds resulting in more uncontrolled violence.

More than a fourth of the acquaintance victims were caregiver victims. The risk of being assaulted or threatened by a patient is sometimes seen as a professional hazard. For example, in a forensic hospital, the one-year incidence rate of physical assault was 70%, and 12% of staff were injured seriously enough to take time off from work (Kelly et al., 2015). In contrast, the number of co-patients who were victimized (30 in total) seems to be relatively small. However, literature has revealed that the prevalence rate of victimization in psychiatric patients is much higher than in the general population, irrespective of the type of sample or type of victimization (Choe, Teplin, & Abram, 2008; de Mooij et al., 2015; Groenhuijsen, 2015; Kamperman et al., 2014). The low number of patient victims in the present study may be due to the usage of court convictions as the data source, whereas other studies used self-reported victimization data (Choe et al., 2008; de Mooij et al., 2015). Again, no interaction effects were found for psychiatric diagnoses, which was in line with another study on MDOs (Goethals et al., 2008).

Clinical implications and further directions

Although violence committed by patients with a major mental disorder is generally a low base-rate phenomenon, preventive action can be taken on several levels since the most likely victims will be acquaintances. First, the present study showed that family members were the most frequent acquaintance victims. This may not come as a surprise since it is estimated that half of the people with a serious mental illness live with their family, and three-fourths of the patients have regular contact with their family (Solomon, Cavanaugh, & Gelles, 2005). In addition, the social network size of MDOs was found to be smaller than that of the general population and to consist mostly of family members (Ter Haar-Pomp, Spreen, Bogaerts, & Volker, 2015). Thus, one explanation for the risk for family members of becoming a victim relates to the fact that they have the most contact with the MDO. Another explanation relates to the finding that there is a poor acceptance of psychiatric pathology among both patients and their family members. As a

consequence, inadequate monitoring and treatment of psychiatric patients often occurs (Marleau, Millaud, & Auclair, 2003; Raymond, Leger, & Lachaux, 2015).

Lewis, Scott, Baranoski, Buchanan, & Griffith (1998) further noted that it was the erosion of the family's capacity to contain the violence that triggered a violent incident rather than a substantive change in the behavior of the patient. Setting limits on the behavior of a sick relative has been identified as another important trigger for violence (Lewis et al., 1998) and is similar to the therapeutic limit setting in institutions. It can be expected that the burden of the care of mentally ill family members will become an even greater burden as family members age, admission periods become shorter, and community-based treatment increases. Family psychoeducation and monitored aftercare should therefore be included in treatment.

Second, caregivers were the second most frequent type of acquaintance victims. The present study did not reveal that a certain type of staff member was victimized more often than others, as was found in previous research (Erdos & Hughes, 2001). Similarly, prior research did not find that some types of care givers were more likely to be violently victimized (e.g., older staff in Decaire, Bedard, Riendeau, and Forrest (2006)). Furthermore, variability in the rates of reporting violence among caregivers is likely to affect the data on the number of victims. A previous study on a similar population found that violent inpatient incidents are rarely prosecuted and adjudicated (author information deleted). As such, official reconviction data are a clear underestimation of the true number of caregiver victims.

It is therefore recommended that careful records of incidents and the related victims be kept to increase awareness of risk factors and situations, inform policy decisions regarding aggression management at the facility level, and eventually prevent more serious future incidents (Kobes, Nijman, & Bulten, 2012). A clear, consistent policy regarding the prosecuting of patients is also recommended (Clark, McInerney, & Brown, 2012; Quanbeck, 2006). At the preventative level, Cornaggia, Beghi, Pavone, and Barale (2011) stressed the need for a "good ward climate" with an appropriate number of nursing staff, a non-overcrowded setting, and adequate staff training (p. 18).

Limitations

A number of restrictions necessitate caution when interpreting and generalizing these findings. First, when using official files to determine recidivism rates, the problem of hidden victimization or *dark numbers* needs to be taken into account. Dark numbers refer to the offenses that are not officially registered. Namely, a large number of offenses are not reported to the police, and the reported offenses do not always result in charges, let alone convictions. As a result, reconviction rates underestimate repeat offenses. This issue might be particularly relevant for family and caregiver victims, who may feel rather reluctant to file a formal complaint. Despite these concerns, we chose to use official recidivism data since they are still considered as a reasonably reliable measure of recidivism (Wartna, 2009).

Second, no comparison group of criminally responsible offenders was used. In a small study, Nestor and Haycock (1997) found that NGRI murderers were more likely to kill blood relatives, especially parents, whereas convicted murderers were more likely to kill a significant other, such as a lover or a spouse. Victims of MDOs convicted of homicide were less likely to be strangers and more likely to be a family member or partner compared to non-MDOs (Shaw et al., 1999). Remarkably, Steadman et al. (1998) found that the targets of violence in a patient group were similar to those of a comparison group of people living in the same neighborhood. Based on these mixed findings, further research with comparison groups is recommended.

CONCLUSION

Knowledge on victim characteristics of NGRI offenders may enhance treatment of offenders and the prevention of future victims. The current study demonstrated that these victims are often adults who are known to the offender and equally likely to be male or female. Further analyses suggested that these profile characteristics of victims are highly similar for perpetrators with different diagnoses. Differences did exist when looking at the different types of perpetrated violence. For example, a gender and age difference emerged when studying sexual assault. Together, our results on the gender, age, and relationship of the victim may inform and help practitioners to recognize potential risk situations.

FOOTNOTES

[1] In Canada, NGRI offenders are referred to as offenders Not Criminally Responsible on Account of Mental Disorder (NCRMD). The term NGRI is used throughout the manuscript referring to similar populations in different countries.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Bader, S., Evans, S. E., & Welsh, E. (2014). Aggression among psychiatric inpatients: The relationship between time, place, victims, and severity ratings. *Journal of the American Psychiatric Nurses Association*, *20*(3), 179-186. doi: 10.1177/1078390314537377

Baxter, H., Duggan, C., Larkin, E., Cordess, C., & Page, K. (2001). Mentally disordered parricide and stranger killers admitted to high-security care. 1: A descriptive comparison. *The Journal of Forensic Psychiatry*, *12*(2), 287–299. doi: 10.1080/09585180122057

Bowers, L., Stewart, D., Papadopoulos, C., Dack, C., Ross, J., Khanom, H., & Jeffery, D. (2011). *Inpatient violence and aggression: A literature review*. London: Kings College.

Choe, J. Y., Teplin, L. A., & Abram, K. M. (2008). Perpetration of violence, violent victimization, and severe mental illness: Balancing public health concerns. *Psychiatric Services*, *59*(2), 153-164. doi: 10.1176/appi.ps.59.2.153

Cirincione, C., Steadman, H. J., & McGreevy, M. A. (1995). Rates of insanity acquittals and the factors associated with successful insanity pleas. *The Bulletin of the American Academy of Psychiatry and the Law*, *23*(3), 399-409.

Clark, C. R., McInerney, B. A., & Brown, I. (2012). The prosecution of psychiatric inpatients: Overcoming the barriers. *The Journal of Forensic Psychiatry & Psychology*, *23*(3), 371-381. doi: 10.1080/14789949.2012.686622

Coid, J., Kahtan, N., Gault, S., Cook, A., & Jarman, B. (2001). Medium secure forensic psychiatry services: Comparison of seven English health regions. *British Journal of Psychiatry*, *178*(1), 55-61.

Cornaggia, C. M., Beghi, M., Pavone, F., & Barale, F. (2011). Aggression in psychiatry wards: A systematic review. *Psychiatry Research*, *189*(1), 10-20. doi:10.1016/j.psychres.2010.12.024

Crocker, A., Seto, M., Nicholls, T., & Côté, G. (2013). *Description and processing of inidviduals found not criminally responsible on account of mental disorder accused of "serious violent offenses"*. Canada: Department of Justice.

Crocker, A. G., Nicholls, T. L., Seto, M. C., Charette, Y., Cote, G., & Caulet, M. (2015). The national trajectory project of individuals found not criminally responsible on account of mental disorder in Canada. Part 2: The people behind the label. *Canadian Journal of Psychiatry*, *60*(3), 106-116. doi: 10.1177/070674371506000305

Daffern, M., Mayer, M., & Martin, T. (2003). A preliminary investigation into patterns of aggression in an Australian forensic psychiatric hospital. *The Journal of Forensic Psychiatry & Psychology*, *14*(1), 67–84. doi: 10.1080/1478994031000074306

Davies, M. (2002). Male sexual assault victims: A selective review of the literature and implications for support services. *Aggression and Violent Behavior*, *7*, 203–214.

Decaire, M. W., Bedard, M., Riendeau, J., & Forrest, R. (2006). Incidents in a psychiatric forensic setting: Association with patient and staff characteristics. *The Canadian Journal of Nursing Research*, *38*(3), 68-80.

Erdos, B. Z., & Hughes, D. H. (2001). Emergency psychiatry: A review of assaults by patients against staff at psychiatric emergency centers. *Psychiatric Services*, *52*(9), 1175-1177. doi: 10.1176/appi.ps.52.9.1175

Estroff, S. E., Swanson, J. W., Lachicotte, W. S., Swartz, M., & Bolduc, M. (1998). Risk reconsidered: Targets of violence in the social networks of people with serious psychiatric disorders. *Social Psychiatry and Psychiatric Epidemiology*, *33*(Suppl 1), S95-101.

Goethals, K., Gaertner, W., Buitelaar, J., & Marle, H. van (2008). Targets of violence and psychosocial problems in psychotic offenders detained under the Dutch Entrustment Act. *The Journal of Forensic Psychiatry & Psychology*, *19*(4), 561-575. doi: 10.1080/14789940801928826

Gow, R., Choo, M., Darjee, R., Gould, S., & Steele, J. (2010). A demographic study of the Orchard Clinic: Scotland's first medium secure unit. *The Journal of Forensic Psychiatry & Psychology*, *21*(1), 139-155. doi: 10.1080/14789940903188956

Gradillas, V., Williams, A., Walsch, E., & Fahy, T. (2007). Do forensic psychiatric inpatient units pose a risk to local communities. *The Journal of Forensic Psychiatry & Psychology*, *18*(2), 261-265. doi: 10.1080/14789940701246360.

Groenhuijsen, M. (2015). Victimological aspects of reporting and not reporting crimes: A general overview. In T. Tollefson (Ed.), *Victimological advances in theory, policies and services* (pp. 111-120). Fresno: Dummond Printing.

Gulayets, M. (2016). Exploring Differences between Successful and Unsuccessful Mental Disorder Defences. Canadian *Journal of Criminology and Criminal Justice*, *58*(2), 1-33. doi: 10.3138/cjccj.2014.E14

Jeandarme, I., Habets, P., Oei, T. I., & Bogaerts, S. (2016). Reconviction and revocation rates in Flanders after medium security treatment. *International Journal of Law and Psychiatry*, *47*, 45-52. doi: 10.1016/j.ijlp.2016.02.033

Johnston, I., & Taylor, P. J. (2003). Mental disorder and serious violence: The victims. *The Journal of Clinical Psychiatry*, *64*(7), 819-824.

Kamperman, A. M., Henrichs, J., Bogaerts, S., Lesaffre, E. M., Wierdsma, A. I., Ghauharali, R. R., . . . Mulder, C. L. (2014). Criminal victimisation in people with severe mental illness: A multi-site prevalence and incidence survey in the Netherlands. *PLoS One*, *9*(3), e91029. doi: 10.1371/journal.pone.0091029.

Kelly, E. L., Subica, A. M., Fulginiti, A., Brekke, J. S., & Novaco, R. W. (2015). A cross-sectional survey of factors related to inpatient assault of staff in a forensic psychiatric hospital. *Journal of Advanced Nursing*, *71*(5), 1110-1122. doi: 10.1111/jan.12609.

Kline, R. (2010). *Principles and practices of structural equation modeling (3rd ed.)*. New York, NY: Guilford Press.

Kobes, M. H., Nijman, H. H., & Bulten, E. B. (2012). Assessing aggressive behavior in forensic psychiatric patients: Validity and clinical utility of combining two instruments. *Archives of Psychiatric Nursng*, *26*(6), 487-494. doi: 10.1016/j.apnu.2012.04.004.

Lewis, M. E., Scott, D. C., Baranoski, M. V., Buchanan, J. A., & Griffith, E. E. (1998). Prototypes of intrafamily homicide and serious assault among insanity acquittees. *The Journal of the American Academy of Psychiatry and the Law*, *26*(1), 37-48.

Liettu, A., Saavala, H., Hakko, H., Rasanen, P., & Joukamaa, M. (2009). Mental disorders of male parricidal offenders: A study of offenders in forensic psychiatric examination in Finland during 1973-2004. *Social Psychiatry and Psychiatric Epidemiology*, *44*(2), 96-103. doi: 10.1007/s00127-008-0419-9.

Livingston, J. D., Wilson, D., Tien, G., & Bond, L. (2003). A follow-up study of persons found not criminally responsible on account of mental disorder in British Columbia. *Canadian Journal of Psychiatry*, *48*(6), 408-415.

Maniglio, R. (2009). Severe mental illness and criminal victimization: A systematic review. *Acta Psychiatrica Scandinavica*, *119*(3), 180-191. doi: 10.1111/j.1600-0447.2008.01300.x.

Marleau, J. D., Millaud, F., & Auclair, N. (2003). A comparison of parricide and attempted parricide: A study of 39 psychotic adults. *International Journal of Law and Psychiatry*, *26*(3), 269-279. doi: 10.1016/S0160-2527(03)00037-2.

Meijwaard, S. C., Kikkert, M., Mooij, L. D. de, Lommerse, N. M., Peen, J., Schoevers, R. A., . . . Dekker, J. J. (2015). Risk of criminal victimisation in outpatients with common mental health disorders. *PLoS One*, *10*(7), e0128508. doi: 10.1371/journal.pone.0128508.

Menezes, S., Oyebode, F., & Haque, M. (2009). Victims of mentally disordered offenders in Zimbabwe, 1980–1990. *The Journal of Forensic Psychiatry & Psychology*, *20*(3), 427-439. doi: 10.1080/14789940802542907

Mezey, G. (2007). Victims and forensic psychiatry: Marginal or mainstream. *Criminal Behaviour and Mental Health*, *17*(3), 131-136. doi: 10.1002/cbm.655.

Miller, J. (1991). Reaction time analysis with outlier exclusion: Bias varies with sample size. *The Quarterly Journal of Experimental Psychology, 43*(4), 907-912.

Mooij, L. D. de, Kikkert, M., Lommerse, N. M., Peen, J., Meijwaard, S. C., Theunissen, J., . . . Dekker, J. J. (2015). Victimisation in adults with severe mental illness: Prevalence and risk factors. *British Journal of Psychiatry*, *207*(6), 515-522. doi: 10.1192/bjp.bp.113.143370

Nestor, P. G., & Haycock, J. (1997). Not guilty by reason of insanity of murder: Clinical and neuropsychological characteristics. *The Journal of the American Academy of Psychiatry and the Law*, *25*(2), 161-171.

Nicholls, T. L., Brink, J., Greaves, C., Lussier, P., & Verdun-Jones, S. (2009). Forensic psychiatric inpatients and aggression: An exploration of incidence, prevalence, severity, and interventions by gender. *International Journal of Law and Psychiatry*, *32*(1), 23-30. doi: 10.1016/j.ijlp.2008.11.007.

Nordstrom, A., & Kullgren, G. (2003a). Do violent offenders with schizophrenia who attack family members differ from those with other victims? *International Journal of Forensic Mental Health*, *2*(2), 195-200. doi: 10.1080/14999013.2003.10471191

Nordstrom, A., & Kullgren, G. (2003b). Victim relations and victim gender in violent crimes committed by offenders with schizophrenia. *Social Psychiatry and Psychiatric Epidemiology*, *38*(6), 326-330. doi: 10.1007/s00127-003-0640-5.

Novak, B., McDermott, B. E., Scott, C. L., & Guillory, S. (2007). Sex offenders and insanity: An examination of 42 individuals found not guilty by reason of insanity. *Journal of the American Acadamy of Psychiatry and the Law*, *35*(4), 444-450.

Parker, I. (1987). Homicide and the insanity defense: A comparison of sane and insane murderers. *Behavioral Sciences & the Law*, *5*(1), 25-35.

Osborne, J. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research & Evaluation, 15*(12), 1-9.

Quanbeck, C. (2006). Forensic psychiatric aspects of inpatient violence. *The Psychiatric Clinics of North America*, *29*(3), 743-760. doi: 10.1016/j.psc.2006.04.011.

Raymond, S., Leger, A. S., & Lachaux, B. (2015). A descriptive and follow-up study of 40 parricidal patients hospitalized in a French secure unit over a 15-year period. *International Journal of Law and Psychiatry*, *41*, 43-49. doi: 10.1016/j.ijlp.2015.03.006.

Shaw, J., Appleby, L., Amos, T., McDonnell, R., Harris, C., McCann, K., . . . Parsons, R. (1999). Mental disorder and clinical care in people convicted of homicide: National clinical survey. *British Medical Journal*, *318*(7193), 1240-1244.

Skipworth, J., Brinded, P., Chaplow, D., & Frampton, C. (2006). Insanity acquittee outcomes in New Zealand. *Australian and New Zealand Journal of Psychiatry*, *40*(11-12), 1003-1009. doi: 10.1111/j.1440-1614.2006.01924.x.

Solomon, P. L., Cavanaugh, M. M., & Gelles, R. J. (2005). Family violence among adults with severe mental illness: A neglected area of research. *Trauma, violence & abuse*, *6*(1), 40-54. doi: 10.1177/1524838004272464.

Steadman, H. J., Mulvey, E. P., Monahan, J., Robbins, P. C., Appelbaum, P. S., Grisso, T., . . . Silver, E. (1998). Violence by people discharged from acute psychiatric inpatient facilities and by others in the same neighborhoods. *Archives of General Psychiatry*, *55*(5), 393-401.

Swanson, J., Swartz, M., Estroff, S., Borum, R., Wagner, R., & Hiday, V. (1998). Psychiatric impairment, social contact, and violent behavior: Evidence from a study of outpatient-committed persons with severe mental disorder. *Social Psychiatry and Psychiatric Epidemiology*, *33*(Suppl 1), S86-94.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed. ed.): New Jersey, NJ: Pearson Education Inc.

Tardiff, K., Marzuk, P. M., Leon, A. C., & Portera, L. (1997). A prospective study of violence by psychiatric patients after hospital discharge. *Psychiatric Services*, *48*(5), 678-681. doi: 10.1176/ps.48.5.678.

Ter Haar-Pomp, L., Spreen, M., Bogaerts, S., & Volker, B. (2015). The personal social networks of personality disordered forensic psychiatric patients. *Journal of Social Work*, *15*(3), 254-276.

Tewksbury, R. (2007). Effects of sexual assaults on men: Physical, mental and sexual consequences. *International Journal of Men's Health*, *6*(1), 22-35. doi: 10.3149/jmh.0601.22

Wartna, B. S. J. (2009). *In de oude fout. Over het meten van recidive en het vaststellen van het succes van strafrechtelijke interventies*. Den Haag: WODC.

Weizmann-henelius, G., & Suutala, H. (2000). Violence in a Finnish forensic psychiatric hospital. *Nordic Journal of Psychiatry*, *54*, 269-273.

General discussion

The research presented in this dissertation was undertaken to provide information about risk factors and recidivism during and after treatment in a medium security unit in Flanders, Belgium. In addition, this dissertation aimed to examine how risk was assessed within a real-world setting what can be described as a daily clinical practice. Twelve research questions were addressed and studied in a large sample of forensic psychiatric patients, who were found not guilty by reason of insanity (NGRI) and subsequently detained under the internment measure and treated in a medium security setting.

MAIN RESEARCH FINDINGS

Judicial system: Comments on some of the criticisms

The Act of July 1, 1964 to 'Protect Society from Abnormals and Habitual Offenders and Offenders who Committed Particular Types of Sex Offenses (Act Protection Society or APS)' underpins the criteria for internment (**Chapter 1**). Offenders who are found not to be criminally responsible *and* offenders who are found to possess diminished responsibility that renders them unable to control their actions are not sentenced but are instead subjected to a protective measure, *internment*. This dichotomized model, which is typical for the Belgian juridical situation in which offenders are declared either criminally responsible or irresponsible, excludes the possibility of combining protective measures and criminal penalties (Vandevelde et al., 2011). After an internment measure is imposed, internees fall under the jurisdiction of one of the regional Commissions of the Protection of Society (CPSs), which will both assess the risk to society and provide the necessary treatment for the internee, with the goal of integration into society. The Internment Act was very progressive for its time and was assessed internationally as a good law, with the exception of the internal and external legal position of the internees, which was assessed as insufficient (Research voor Beleid, 1995). However, several criticisms have been raised regarding the internment measure (e.g., Casselman, Devuysere, & Vervaeke, 2003; De Clercq & Vander Laenen, 2013). Some of these criticisms were addressed in this dissertation.

In **Chapter 2**, it was found that in comparison to other countries, Belgium is the only country that does not require a psychiatric assessment to impose mandatory treatment. In addition, criteria for both experts and criminal responsibility evaluations are lacking, whereas

examples for good practices can be found in other countries. The poor quality and lack of criteria concerning both experts and expert reports have been noted by several authors (De Clercq & Vander Laenen, 2013; Vandevelde et al., 2011) and became evident when collecting the data in the current research. The quality of the expert reports appeared to vary substantially. Other problems related to criminal responsibility evaluations were characterized by a lack of multidisciplinarity, a lack of clinical observation and insufficient financial compensation, with the latter being often cited as a reason for the shortage of experts. On the other hand, it was found that most countries struggle with the difficult task of determining a legal standard for insanity (**Chapter 2**).

A second criticism raised by several authors relates to the lack of a common policy across different CPSs during the execution phase (Casselman et al., 2003; Delva, 1999; Goethals, 1985; De Ruyver & Goethals, 1991). The Internment Commission was very critical regarding the diversity and the stereotyped nature of different CPS decisions, arguing that the composition of the CPS was unbalanced and that there was a lack of training and continuing education (Delva, 1999). The current study could not entirely confirm this assumption. For example, although there was a notable shortage of clinical information made available to the CPSs at every step of the execution of the internment measure, the CPSs managed to make rational release recommendations. As is shown in **Chapter 10**, the internees who were granted an unconditional release from the internment measure were internees for whom the risk of reoffending could be regarded as low based on their criminogenic needs, in accordance with the Risk Need Responsivity (RNR) principles (Bonta & Andrews, 2007). In other words, CPS members based their decisions on those risk factors that are empirically proven to be related to general and violent recidivism. In addition, no difference was found among CPSs with respect to recidivism and revocation rates. Thus, the results appear to contradict some of the abovementioned criticisms. This finding also stands in contrast to a study that investigated conditional release evaluations in the United States. This study showed not only that there was a lack of uniformity concerning any aspect of the decision-making process but also that the evaluators de-emphasized important factors that are empirically related to recidivism (Gowensmith, Bryant, & Vitacco, 2014). In Flanders, Vandenbroucke (1981) also questioned the conservative release policy, as he claimed that internees were often revoked for minor violations. This topic was investigated in **Chapter 5**.

Interestingly, this study shows that more than 40% of the decisions to revoke conditional release were made after a non-crime-related incident (e.g., absconding and non-compliance) was reported to the CPSs. These type of incidents can be regarded as rather minor violations and thus confirm the former statement. However, it should be noted that relatively minor incidents may reflect a long journey that involves more serious incidents and therapy-interfering behavior, leading to the hospital's decision to stop treatment. Taken together, the criticisms raised about the CPSs appear to be overly pessimistic.

Finally, more specifically, the administrative internment (*Art. 21 procedure)* was heavily criticized on legal grounds. This article holds that in addition to regular internment measures, the Minister of Justice can impose an internment measure on a convict when, during detention, the convict is found to be "in a state of insanity, a serious state of mental disturbance or mental deficiency, which renders him/her unfit to control his/her actions" (Art. 21 APS). The procedure takes place without an adversarial debate. Heavy criticism regarding the weak legal position of the convicted internee led to the proposal to abolish the procedure in the newly proposed law. In **Chapter 4**, the implications of this proposal for the regular mental health care system are discussed. It is shown that convicted internees suffer from severe psychiatric problems similar to other internees and also represent a higher risk for recidivism. Generally, these findings were in line with previous research on convicted internees in Wallonia (Vicenzutto & Pham, 2015). It became clear that abolishing the Art. 21 procedure would resolve a number of legal issues related to legal uncertainty but would create a serious challenge to regular psychiatric hospitals, especially in an era of de-institutionalization and budget cuts. Regular psychiatric units currently do not have the required infrastructural and relational security levels, nor do they have the skills required to deal with this population. At the hearing commission regarding the reparation of the new law proposal, these research findings were discussed. Ultimately, it was decided to keep administrative internment in the new law.

The medium security patient: Complex profile

In **Chapter 3**, it is shown that the forensic population admitted to a medium security unit comprises a complex population with many psychopathological, social and criminological needs.

Typically, the Flemish internee can be described as a low-educated young man of Belgian nationality who is diagnosed with severe psychiatric disorders and comorbid diseases. His past work and relationships are unstable, and he usually lives alone, having no partner or children at the time of the index offense. The index offense is often a violent offense, in which several adult individuals who are known to the internee (often a family member) are victimized (**Chapter 12**). These sociodemographic characteristics and index offenses are very similar to those reported in international research (Blattner & Dolan, 2009; Coid, Kahtan, Gault, Cook, & Jarman, 2001; Crocker, Nicholls, Seto, & Côté, 2015; Dolan & Khawaja, 2004; Gow, Choo, Darjee, Gould, & Steele, 2010; Melzer et al., 2004). However, differences were found with respect to clinical diagnoses, with a greater number of personality disorders and substance misuse found in the Flemish population, as well as a smaller number of psychotic disorders (Blattner & Dolan, 2009; Carr et al., 2006; Dolan & Khawaja, 2004; Gow et al., 2010; Gradillas, Williams, Walsh, & Fahy, 2007). In addition, the criminal history of the internees was more pronounced (e.g., 4.6 years of detention prior to the first admission to a medium security facility) and the risk level according to the HCR-20 was higher for the Flemish population than reported in international research (Blattner & Dolan, 2009; Carr et al., 2006; Dolan & Khawaja, 2004; Freestone et al., 2012; Gow et al., 2010; Gradillas et al., 2007; Gray et al., 2004; Gray, Taylor, & Snowden, 2008; Hildebrand & de Ruiter, 2004). In this respect, internees were more comparable to Dutch patients detained under the Entrustment Act (Mudde, Nijman, van der Hulst, & van den Bout, 2011; de Vogel & de Ruiter, 2006). These findings may partly explain why release into the community occurred less frequently than reported in other studies (Blattner & Dolan, 2009; Dolan & Khawaja, 2004; Gow et al., 2010) and why readmissions to medium security occurred more often than reported in other studies (e.g., Coid et al., 2001).

Another interesting finding was the high percentage (82%) of internees with prior admissions to general psychiatric services. Although these results were in line with previous research (Blattner & Dolan, 2009; Gow et al., 2010; Linhorst & Scott, 2004; Melzer et al., 2004), they question the role of general mental health services. For example, Brand, Mellsop, and Tapsell (2015) examined psychiatric care provided in the year prior to the offense and found that access to care was not the problem. Instead, a non-assertive approach to treatment was the

main problem. During data collection, similar examples were found in the researched files; however, this issue was not investigated further in the present dissertation.

Substantial treatment drop-out

In terms of treatment characteristics, important drop-out rates were identified. One-third of the study population dropped out prematurely from medium security treatment, and 41% had intermittent treatment courses due to one or more time-out periods. Drop-out from treatment resulted in lengthy re-imprisonment periods, lasting on average 1.8 years (**Chapter 5**).

Preventing and avoiding drop-out from treatment is very important for several reasons. First there is a higher risk of recidivism in cases of treatment non-completion (McMurran & Theodosi, 2007; Olver & Wong, 2009). Furthermore, the chance of further treatment in another setting may be compromised. Carr et al. (2006) explained the high rate of forensic treatment drop-out by citing poor treatment preparation during detention. As shown in **Chapter 3**, detention periods prior to the first medium security treatment were extensive (on average, 4.6 years of detention). Research revealed that inmates adopt attitudes that are adaptive in correctional settings (such as distrust towards staff, intimidating behavior and dissimulation of symptoms) but become maladaptive once released (Rotter, Carr, Magyar, & Rosenfeld, 2011). Moreover, inmates with a mental disorder tend to have even more difficulty coping with this adaptation (Carr et al., 2006). In view of the lack of sufficient treatment possibilities in Belgian prisons, pre-therapeutic counseling should be started in prison to reduce the adaptation process (or the culture shock) that mentally ill offenders undergo after transfer from a prison to a treatment setting (Carr et al., 2006). This type of system has been partly developed in recent years in Flanders (Stassen, Habets, Mertens, De Laender, & Jeandarme, 2014).

Treatment drop-out also raises the question of whether the medium security level that was assigned to these patients on a clinical basis was adequately matched to their actual risk and security level. Understandably, in an effort to get as many internees out of prison as possible, medium security units may have been tempted to accept patients with a high risk and security level. During the study period, there were no high security units in Flanders, and, as argued by Coid & Kathan (2000), the absence of alternatives can play a role in admission criteria. Another hypothesis is that treatment programs were not responsive enough to offender characteristics,

as mentioned in the RNR model by Bonta & Andrews (2007). For instance, a large number of intellectually disabled patients and a sizable minority of psychopathic patients were identified, even though the MSUs do not provide treatment for these particular problems. In **Chapter 7**, it is shown that internees with a high score on the Psychopathy Checklist-Revised (PCL-R; Hare, 2003) were a particularly problematic group that engaged in more therapy-interfering behavior. In addition, drug-related incidents during medium security treatment were found in more than half of the population. Together with the high percentage of substance misuse diagnoses (**Chapter 3**), this finding highlighted the importance of focusing on ongoing substance misuse treatment or access to addiction services. It is likely that these treatments were not fully available at the time the units started.

In addition to hospital-initiated expulsion, it is interesting to note that some of the re-incarcerations (11% of the revocations) were requested by the patients themselves (**Chapter 5**). Internees residing in medium security units may feel too much pressure in treatment and a lack of control in comparison to their time in prison (To, 2015). Preliminary results from another ongoing study confirmed that internees in a medium security setting perceive more stress, especially when staff is present (Habets, Delespaul, & Jeandarme, 2016). Nonetheless, establishing a positive working alliance with the patient is an important prerequisite for the rehabilitation of offenders so that criminogenic factors can be controlled and the risk of reoffending can be reduced or prevented. Further research on how to establish a treatment program that is responsive to the forensic population is recommended (Gannon & Ward, 2014). In addition, maintaining careful records of incidents can increase awareness of risk factors and situations, inform policy decisions regarding aggression management on the facility level and eventually prevent more serious incidents and revocations in the future.

Low post-treatment recidivism rates

In **Chapter 10**, recidivism data based on criminal records show that recidivism rates were reassuringly low (i.e., only 12.5% of the internees reoffended with a general offense and 7.4% with a violent offense). These findings are equivalent to or even lower than those reported in other studies, which found rates of 7.1% to 63% for general recidivism and 1.8% to 46.0% for

violent recidivism (e.g., Hayes, Kemp, Large, & Nielssen, 2014). In addition, in line with Hayes et al. (2014), none of the unconditionally released patients reoffended in a serious crime.

Preventing future criminal behavior is the most important goal of forensic psychiatric treatment (Menghini, Ducro, & Pham, 2005). Therefore, these findings are important. The significant decrease in offenses between the pre-treatment and post-treatment periods suggests that forensic psychiatric treatment is effective. However, without a randomized control trial, it remains unclear whether these changes can be attributed to the effect of treatment. The study also showed that the risk-reducing effect was most apparent when the interment measure was in effect (i.e., the patient was under some form of supervision). Key components of the conditional release process for NGRI patients include the development and monitoring of conditions of release and the possibility of revocation and inpatient hospitalization when violations of conditions occur (Dirks-Linhorst & Linhorst, 2006). In the context of risk management and the prevention of recidivism, most conditionally released individuals are required to undergo treatment and (probation) supervision. Not adhering to prescribed rules and ancillary conditions often results in a return to a secure inpatient facility for further treatment and/or confinement. Therefore, in NGRI populations, two outcome metrics related to *failure* are typically being used: new criminal charges and conditional release revocation due to criminal acts and/or rule violations. In **Chapter 10**, it is shown that similar to other studies, revocations for rule violations were more frequent than revocations for the acquisition of new criminal charges (Vitacco, Vauter, Erickson, & Ragatz, 2014; Wiederanders, 1992). According to Wiederanders, Bromley, and Choate (1997), revocations can be identified as a major cause of low re-offense rates. However, the relationship between revocations and re-offenses was not linear.

Given the fact that revocation can occur as a result of a new criminal act that is not adjudicated, only using recidivism rates based on the Central Criminal Record may underestimate the real rate of recidivism. In **Chapter 11**, it is shown that significantly more internees had incidents reported to the CPS than official recidivism, with a five-fold increase in both general and violent relapse. This finding underscores the need to be very specific when comparing outcome metrics for NGRI offenders, as different types of recidivism have been studied, such as reconvictions, re-arrest, revocation and (re)incarceration rates and self-reports (Heilbrun & Griffin, 1993). The rather significant base rate increases observed after the inclusion of incidents

indicates the importance of the choice of outcome measure. Future research should preferably make use of various outcome measures, including police arrest data if available.

Risk factors for inpatient and outpatient violence

Risk factors for violence were assessed both during and after medium security treatment. Not only do risk factors differ when conducting research in either in- or outpatient samples, studying inpatient violence separately is important for several other reasons. First, inpatient violence affects the stability of an institution, impacts staff turnover and has a negative impact on the therapeutic process (Gow et al., 2010; Quanbeck, 2006). Besides affecting staff, patients perpetrating violence are also adversely affected, as seclusion and restraints are often used to manage aggressive behavior and inpatient violence can result in prolonged involuntary confinement. Inpatient violence has also been associated with recidivism after treatment. In **Chapter 6**, it is shown that violent incidents can cause a patient to be expelled from treatment, thus further contributing to an increased risk of recidivism.

Contrary to the hypotheses developed based on a literature review of the risk factors for violence, few differences were found between internees who relapsed with a violent offense and internees who did not (**Chapter 6**). It should be noted that the occurrence of violence involves an interplay between individual, situational and structural factors. Therefore, further research should adopt a more dynamic or contextual approach; in the present research, only individual risk factors could be studied.

Regarding risk factors for recidivism after treatment, it is shown in **Chapter 11** that past behavior predicts future behavior (i.e., more violent recidivists were identified as habitual offenders). In this particular study, which was conducted on a subpopulation, the low base rate of recidivism and the relatively small number of participants – and hence the loss of statistical power – could be an explanation for the limited results that were found.

Regarding risk factors for inpatient violence, it is shown in **Chapter 6** that interpersonal violence is associated with general misconduct during treatment (i.e., more violent patients absconded, did not comply with hospital rules, and kept drinking alcohol during treatment even though this was prohibited). In fact, after controlling for other variables, treatment characteristics were the only characteristics that were predictive of interpersonal violence, whereas the

commonly used risk assessment instruments and other well-established risk factors for violence, such as previous violence, were not found to be significant predictors. Taken together, these results suggest that dynamic treatment variables can act as warning signs for interpersonal violence. Furthermore, monitoring treatment processes and inpatient incidents, preferably using standardized instruments, is important. In this study, the small number of patients involved in verbal and physical violence was remarkable, particularly given the fact that most of the patients had a history of violent (sexual) convictions. One explanation could be that this low rate of violence was the result of adequate risk management on behalf of the treatment staff. It is generally assumed that forensic psychiatric patients bear a great risk for inpatient violence in comparison to general psychiatric settings (Bowers et al., 2011; Linhorst & Scott, 2004). However, when only physical aggression was considered, the differences among acute, general and forensic psychiatric settings were no longer significant (Bowers et al., 2011). Similarly, after controlling for demographic and clinical variables, Linhorst and Scott (2004) found that non-forensic patients treated during the same period in the same hospital were just as likely as forensic patients to exhibit violent behavior.

On a general note, studies of violent incidents also revealed that the definition of a violent (or aggressive) incident in the literature should be carefully scrutinized because in addition to physical violence towards others, verbal violence and/or violence towards oneself or objects can also be included (Alia-Klein, O'Rourke, Goldstein, & Malaspina, 2007; Cullen et al., 2015; Daffern, Duggan, Huband, & Thomas, 2008; Decaire, Bedard, Riendeau, & Forrest, 2006; Gow et al., 2010; Gudjonsson, Rabe-Hesketh, & Wilson, 2000). Unfortunately, the prevalence of these different types of violence is not always separately described separately in the literature (Daffern et al., 2008; Decaire et al., 2006).

Reporting incidents to the legal authorities

To our knowledge, empirical studies that focus specifically on the prevalence of reporting violent and/or non-violent (e.g., theft or drug-related offenses) crime-related incidents to police or judicial authorities during forensic psychiatric treatment are quasi non-existent (Clark, McInerney, & Brown, 2012).

In **Chapter 5**, it is shown that crime-related incidents often fail to lead to new convictions (4.9%). This finding is in accordance with research in Dutch general psychiatry, which revealed that only 10% of the physically violent incidents reported to the police were brought to court (Harte, Van Leeuwen, & Theuws, 2013). A study in New Zealand suggested that the referral of violent patients for prosecution remains rare and arbitrary, despite being increasingly mentioned as an option for staff (Kumar, Fischer, Ng, Clarke, & Robinson, 2006). To our knowledge, no studies investigated whether prosecution and conviction rates in forensic settings differ significantly from those in general psychiatric settings, but further research should focus on this matter. One study in a high security hospital in England found that successful prosecutions do occur in forensic settings, although at a low rate (e.g., 41 successful prosecutions for assaults on staff or slightly over 10% of total staff assaults) (Clark et al., 2012).

A low reconviction rate can be the consequence of a low rate of reporting, prosecuting or sentencing. As shown in previous research, the decision to report for example a violent incident to police services can be influenced by different factors, including the influence of the institute's policy and the attitudes of colleagues (Harte et al., 2013). In addition, treating psychiatrists are likely to be confused concerning conflicts of interest and confidentiality issues. Even if an incident is reported, police enquiries might be hampered by limited cooperation both by patients due to their psychiatric symptoms and by clinicians due to confidentiality issues. Further, prosecuting psychiatric inpatients is often seen as a poor use of limited resources, and it could be argued that the protection of society is already accomplished by hospitalization, while the risk for other patients and staff is perceived as being "part of the job" (Dinwiddie & Briska, 2004). Unfortunately, our research did not have access to data related to police enquiries or prosecution practices. In addition, in the medium security units, there was no clear policy related to the reporting of crime-related incidents. The results did show that crime-related incidents were reported – mostly indirectly through probation officers – to the supervising CPS, but it was not known if this reporting resulted in a formal complaint. Information was also available on the sentencing level, where it became obvious that few crime-related incidents were prosecuted and subsequently adjudicated (**Chapter 5**). Most crime-related incidents were less serious in nature, which might be a reason to renounce prosecution. Furthermore, in the specific case of the Belgian internment measure, the lack of prosecution may also be due to the mandatory

supervision of the CPS, which allows the prosecutor to immediately re-incarcerate the internee via a rather flexible procedure without contradictory debate. This way, the lengthy and lingering procedure of a new trial, which would in most cases result in yet another internment measure, is avoided. Clark et al. (2012) showed that a consistent approach within a local agreement between psychiatrists and the investigating and prosecuting authorities can overcome some of the abovementioned barriers. Therefore, a clear, consistent policy regarding the prosecution of patients is recommended (Clark et al., 2012; Quanbeck, 2006).

Scant research on risk assessment field validity

Empirically based risk assessment using an actuarial or a structured professional judgment (SPJ) approach has become a part of routine practice in forensic psychiatry (Fazel, Singh, Doll, & Grann, 2012). Surprisingly, the field reliability and validity of risk assessment instruments have never been tested in Belgium. Even within a European context, data are very scarce. **Chapters 8 and 9** show the relevance of studying clinical risk assessment within a naturalistic setting.

In **Chapter 8**, it is shown that the Historical Clinical Risk Management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997) was not accurate in identifying high-risk individuals, as demonstrated by a low positive predictive value (PPV). The negative predictive value (NPV) was higher than the PPV, indicating that the HCR-20 results were more accurate in making 'rule out' decisions (i.e., identifying individuals who are at low risk and ready for discharge). This finding is consistent with the suggestion of Fazel et al. (2012) to use risk assessment tools to screen out low-risk cases. The predictive validity analyses further showed that AUCs were non-significant and lower than those found in research designs in general (Pham, Ducro, Marghem, & Réveillère, 2005; Singh, Grann, & Fazel, 2011). In comparison to other field validity studies, the AUCs were similar to those reported in the study of Vojt, Thomson, and Marshall (2013) but lower than those reported in the study of Pedersen, Ramussen, and Elsass (2012). More specifically, only a small number of items showed predictive validity. An interesting finding, which is in line with previous research (Coid et al., 2011; Mudde et al., 2011), was that the item *Impulsivity* was predictive both during and after treatment and thus appears to be an interesting dynamic variable on which to focus treatment (Coid et al., 2011; Mudde et al., 2011). It should be noted that this variable was

proposed by Vertommen and Maesschalck (2007) as one of the factors that could be used to identify a medium security patient (**Chapter 3**).

In **Chapter 9,** it is shown that the field validity of the Psychopathy Checklist Revised (PCL-R; Hare, 2003) is equally poor. The PCL-R is an extensively used and researched instrument for assessing psychopathy and is also frequently introduced in the legal arena to inform violence risk assessment (DeMatteo et al., 2014), either in isolation or as an important component within risk assessment instruments, such as the Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice, & Cormier, 2006) and the HCR-20 (Webster et al., 1997). The results of this field validity study revealed interrater disagreement well beyond what would be expected based on the reliability data reported in the PCL-R manual (Hare, 2003) and add to the growing evidence for poor intraclass correlation coefficients (ICCs) in field settings (DeMatteo et al., 2014; Edens, Cox, Smith, DeMatteo, & Sorman, 2015; Lloyd, Clark, & Forth, 2010; Murrie, Boccaccini, Johnson, & Janke, 2008; Murrie et al., 2009). When the repeated measures were explored both across (hospital vs. prison) and within settings, substantial differences were found, with only a minority of the difference scores falling within 1 standard error of measurement (SEM). Although differences of more than six points – or two SEMs – should be extremely rare, such differences were found for almost half of the scores. Our findings are in line with other recent field validity studies that have reported relatively high percentages of cases with differences exceeding two SEMs (Edens et al., 2015; Murrie et al., 2008; Sturup et al., 2014). Following Fleiss' (1986) criteria for interpreting reliability statistics, only Facet 4 ($ICC_{A,1}$ = .59–.65) would be considered 'good,' which was also in line with other field validity studies (Miller, Kimonis, Otto, Kline, & Wasserman, 2012; Sturup et al., 2014), which found the highest ICCs for Facet 4 scores. Research that investigated rater differences in applied settings has focused primarily on two non-mutually exclusive explanations for these interrater differences: adversarial allegiance and individual differences across raters. Although not adversarial allegiance per se, we expected that the current research would show evidence of some type of partisanship or contextual pressures that might impact scores across these settings. For example, given systemic pressures within the prison system, it was argued that prison evaluators might have felt compelled to score an examinee lower on the PCL-R than in forensic institutions in order to facilitate the transfer of forensic patients from prison to hospital. Some evidence for this hypothesis was found. Namely, when comparing repeated

measures for the same offender across settings, the mean prison scores were significantly lower than the mean hospital scores. However, the analyses were based on small numbers and warrant further research.

Although poor IRR does not necessarily indicate that the predictive validity will be low, this is generally what was found. The predictive validity was poor, especially for the PCL-R total score and for Factor 1, which did not predict general or violent recidivism. Factor 2 scores significantly predicted general and violent recidivism. Surprisingly, on the facet level, Facet 3 scores were the only significant predictors of general and violent recidivism. These findings were somewhat surprising given the extensive research demonstrating the good predictive validity of the PCL-R with regard to general and violent recidivism in well-controlled research designs (Hemphill, Hare, & Wong, 1998; Leistico, Salekin, DeCoster, & Rogers, 2008; Salekin, Rogers, & Sewell, 1996; Walters, 2003; Yang, Wong, & Coid, 2010), although the results for Factor 1 in recent meta-analyses have been quite poor. However, our study results were in line with prior research that identified null findings with respect to predictive validity in field settings (Murrie, Boccaccini, Caperton, & Rufino, 2012; Neal, Miller, & Shealy, 2015). When comparing AUCs for the repeated measures across settings, no significant difference in AUCs was found, although the difference between the prediction of violent recidivism nearly reached significance in favor of the hospital scores (AUC = .66 vs. AUC = .35, p = .05). However, it should be noted that these analyses were based on small numbers. At this point, no firm conclusions can be drawn from the current research regarding the superiority of different groups of raters in terms of predictive validity.

While the PCL-R leaves some room for subjectivity in scoring, the level of discrepancies found in the current study should raise serious concerns, particularly when considered in conjunction with the poor predictive validity results obtained for the total score and Factor 1. The potential adverse implications of these field validity findings should not be taken lightly, given concerns about the stigmatizing effects of the 'psychopath' label (e.g., Bersoff, 2002; Edens, Davis, Fernandez Smith, & Guy, 2013; Lloyd et al., 2010; Wayland & O'Brien) and particularly Factor 1-type characteristics (Edens, Colwell, Desforges, & Fernandez, 2005; Edens et al., 2013; Sundby, 1997) on legal decision-making.

The questions of why field validity is poor and what can be done to improve field validity remain. Potential explanations relate to reliability issues, such as negligence or insufficient

experience in scoring. For example, it is argued **in Chapter 9** that raters who do not frequently use the PCL-R should not use the instrument at all because scores from prolific raters tend to outperform scores from less prolific raters (Murrie et al., 2012). It might be advisable to have few evaluators doing all the ratings in settings where the PCL-R is not used on a regular basis. An interesting question concerning which raters are best fit for the job emerged during the research. Should risk assessment be performed by psychologists (with or without a doctoral degree), psychiatrists or criminologists? Unfortunately, no data were available regarding the individual staff members who performed the scoring using the PCL-R and/or the HCR-20 (only place of administration); therefore, no analyses could be conducted to investigate issues such as the effects of the level and type of training, experience, prolific raters, the method of assessment (file versus interview) or the personality traits of the evaluators. The question therefore remains unanswered and is a topic for further debate. Meanwhile, it increasingly seems advisable that risk assessment should be performed by two independent, trained and experienced raters, ideally by using a consensus score, as proposed by de Vogel and de Ruiter (2006). In addition, after following initial training, continuous peer review processes can minimize drift from item descriptions. In practice, however, it is often not possible (due to financial or time restraints) to adhere to all of these recommendations. Another possible explanation relating to poor field validity was discussed in **Chapter 8** and relates to base rate neglect, which is a well-recognized problem in the risk assessment literature. It is therefore important that recidivism data are now available to clinicians (**Chapter 10**). However, research has shown that this issue may remain problematic in clinical practice, as only one out of five raters makes effective use of this information (Walters, Kroner, & DeMatteo, 2014). A final consideration related to the poor predictive validity is that the assessment and thus the knowledge of a risk score and high risk status resulted in more effective risk management, which would obviously attenuate effect sizes. Ultimately, the usefulness of risk assessment can be judged by its ability to contribute to harm reduction. While this explanation is tempting, more in-depth research is needed to prove this assumption. Although research has focused almost exclusively on predictive validity, it remains unclear whether the use of tools for structured professional judgment actually helps prevent crimes (Wand, 2012). Based on a randomized controlled study that investigated this research question in the Netherlands using the Short Term Assessment of Risk and Treatability (START;

Webster, Martin, Brink, Nicholls, & Desmarais, 2009), it was concluded that the goal of risk prevention was not achieved (Troquete et al., 2013). In sum, field validity studies, such as the current studies, are important for clinicians to consider when conducting assessments in practice. Such studies are also important for researchers to consider when developing and refining new instruments (Neal et al., 2015). For instance, in **Chapter 8**, it is argued that a drawback in the literature is that nearly all studies of the predictive validity of risk assessment tools report AUC values, as an outcome measure. This parameter indicates the probability of a randomly selected recidivist having a higher risk classification than a randomly selected non-recidivist. However, this retrospective measure does not provide a full picture of the predictive value of a risk assessment tool. In accordance with Singh (2013), it is therefore recommended to include more prospectively orientated statistics to address how well a tool's prediction of risk agrees with the actual observed risk (PPV, NPV, number needed to detain (NND), number safely discharged (NSD)). This approach might provide more useful information, as such statistics simulate clinical decision-making.

MAIN LIMITATIONS AND STRENGTHS

Several limitations of the present research deserve attention. First, this multicenter, retrospective study was based on case note material obtained for clinical rather than research purposes. Therefore, experimental manipulation and systematic controls could not be conducted. Not only were several hospitals involved, but there were also several clinicians, and the recording systems differed in rigor. This approach resulted in missing data, which were very evident when examining the risk assessment instruments or the intelligence scores.

A second limitation relates to the problems that arise when comparing study results to international findings. Caution is warranted due to differences in methodology, major differences between the relevant legal systems and the organization of forensic health care, and variations in the the characteristics of local patient groups and local treatment providers (Melzer et al., 2004; Salize & Dressing, 2007). Different studies use significantly different patient samples, different definitions of incidents and recidivism and different assessments of incidents, which may cause problems for comparability and generalizability. For example, in this dissertation, incident

reports were collected retrospectively on a clinical basis, not using a standardized instrument such as the Staff Observation Aggression Scale-Revised (SOAS-R; Nijman et al., 1999; Palmstierna & Wistedt, 1987). As such, methodological issues hamper the comparability of the current results with other studies.

Finally, the scope of this dissertation was limited in that it examined admissions retrospectively. For example, the study did not provide insight concerning how incidents influenced the therapeutic relationship between patients and hospital staff (Cornaggia, Beghi, Pavone, & Barale, 2011). In addition, data from the three forensic units were not presented separately, so possible differences in treatment approaches could not be analyzed.

While the limitations of the dissertation are important, so are its strengths. First, the study covered virtually all (98%, *N* = 531) medium security admissions over an extensive period in Flanders. It is important to note that this population may not be representative of the Flemish interned population as a whole. Nevertheless, this population forms an important group, first because of its complex risk and need profile and second because treatment may be most effective in this group (Skeem, Manchak, & Peterson, 2011). In light of the paucity of research on the interned population in Flanders in general, this research can be considered to be a big step forward.

Second, extensive effort was made to cross-reference all case note material by using multiple sources of information. For example, inconsistencies were found regarding the clinical diagnoses, which were adjusted after consulting the treating psychiatrists in nearly half of all cases.

Finally, the study enabled research on field validity, which is a rather neglected but highly relevant topic in the literature.

SOME CONCLUDING REMARKS AND RECOMMENDATIONS FOR FURTHER RESEARCH

After years of criticism concerning the internment system and several convictions by the European Court of Human Rights (resulting in a landmark decision on September 6, 2016) for keeping internees in prison without appropriate psychiatric treatment, there has been growing interest in this system from politicians, clinicians, and the public. However, despite the ongoing

debate regarding internment and internees, nationwide, there is still a surprising shortage of basic information on this population, as well as a lack of outcome monitoring for categorical forensic psychiatric services. This dissertation provided some in-depth information about one category of internees, namely those who require a medium risk and security level. More research is needed to provide a complete picture of this population, especially regarding an understudied group: those receiving ambulatory care and those hospitalized in general psychiatric services. There is currently no nationwide systematic collection of data regarding population demographics, incidents or recidivism rates. Information is fragmented across institutions, resulting in missing information. My first recommendation would therefore be the nationwide systematic collection of basic data regarding population characteristics and recidivism rates.

Next, it is paramount to further investigate which interventions are effective for which internees and in which settings such interventions should be implemented. Categorical forensic care is expensive and should only be instigated when needed, but in order to do so, the security need should be adequately assessed. At this point, there is room for improvement. Identifying and constructing distinct patient profiles may be an interesting first step in this direction (van der Veeken, Bogaerts, & Lucieer, 2015). In addition, clear directions concerning when and how to transfer NGRI offenders to general psychiatric units, which typically do not adhere to the RNR principles (Bonta & Andrews, 2007), are needed. However, recent initiatives in Flanders appear to be concerned primarily with getting more internees out of prison as soon as possible and less concerned with the adequate matching of the treatment setting to the risk level (Jeandarme, 2016). Getting people out of prison is one thing, but keeping people out of prison is another. My second recommendation is to pay more attention to the systematic allocation of patients to appropriate security levels because this issue is central to the operation of forensic mental health services.

Forensic psychiatric care has made important advances since the 1970s, when Martinson stated that few programs were available that could successfully reduce violent recidivism. Despite these achievements, there is still little research on NGRI offenders, with no meta-analysis of treatment outcomes available at this point. The current study adds to this scant research and provides some evidence for continuing to use forensic care, yet it also found that there is room for improvement, for instance with respect to treatment drop-out rates and the application of

risk assessment tools. My third recommendation is to further study the risk factors for treatment drop-out and focus on therapeutic relationships.

My fourth recommendation is to continue the study of risk assessment implementation. Assessing risk on a systematic basis is considered to be an important aspect when following the RNR-model (Bonta & Andrews, 2007). The results obtained in the present dissertation regarding field validity have shown that raters should be aware that the utility of their assessments (and the therapeutic and judicial decisions that stem from it) will drop drastically when the abovementioned recommendations are not followed.

Finally, at the time of the writing of this dissertation, the APS was still in effect. However, the new internment law will be implemented shortly, and some major changes will be implemented. For example, internment will no longer be possible for minor crimes, and the procedure to revoke a patient on conditional release will become more difficult. The current system, which enables a very close working relationship between judicial authorities and clinicians, was identified as one of the key success elements of the internment system in **Chapter 10**. The consequences of introducing a more legal rights-based model with additional bureaucratic procedures and due process should therefore be carefully scrutinized. On the other hand, a decline in the number of re-incarcerations can be foreseen. In light of the lack of treatment possibilities for mentally ill offenders in Belgian prisons, this change would be a big step forward. At the same time, such a decline might create new problems because institutions will face more unmotivated, difficult-to-treat and violence-prone patients. My final recommendation is to study and evaluate the (criminal) outcomes of the upcoming law reform.

REFERENCES

Alia-Klein, N., O'Rourke, T. M., Goldstein, R. Z., & Malaspina, D. (2007). Insight into illness and adherence to psychotropic medications are separately associated with violence severity in a forensic sample. *Aggressive Behavior, 33*(1), 86-96. doi: 10.1002/ab.20170

Bersoff, D. N. (2002). Some contrarian concerns about law, psychology, and public policy. *Law and Human Behavior, 26*(5), 565-574.

Blattner, R., & Dolan, M. (2009). Outcome of high security patients admitted to a medium secure unit: The Edenfield Centre study. *Medicine, Science and the Law, 49*(4), 247-256. doi: 10.1258/rsmmsl.49.4.247

Bonta, J., & Andrews, D. A. (2007). *Risk-need-responsivity model for offender assessment and treatment (User Report No. 2007-06)*. Ottawa, ON: Public Safety Canada.

Bowers, L., Stewart, D., Papadopoulos, C., Dack, C., Ross, J., Khanom, H., & Jeffery, D. (2011). *Inpatient violence and aggression: A literature review*. London: King's College London.

Brand, E., Mellsop, G., & Tapsell, R. (2015). A review of the psychiatric care provided to patients who subsequently offended. *International Journal of Clinical Psychiatry and Mental Health, 3*, 31-35. doi: 10.12970/2310-8231.2015.03.01.5

Carr, W. A., Rotter, M., Steinbacher, M., Green, D., Dole, T., Garcia-Mansilla, A., . . . Rosenfeld, B. (2006). Structured Assessment of Correctional Adaptation (SACA): A measure of the impact of incarceration on the mentally ill in a therapeutic setting. *International Journal of Offender Therapy and Comparative Criminology, 50*(5), 570-581. doi: 10.1177/0306624X06289176

Casselman, J., Vuysere, S. De, & Vervaeke, G. (2003). De Commissies ter Bescherming van de Maatschappij van Antwerpen, Gent en Leuven. Een vergelijking van de geïnterneerdenpopulatie. *Panopticon, 6*, 604-609.

Clark, C. R., McInerney, B. A., & Brown, I. (2012). The prosecution of psychiatric inpatients: Overcoming the barriers. *The Journal of Forensic Psychiatry & Psychology, 23*(3), 371-381. doi: 10.1080/14789949.2012.686622

Clercq, M. De, & Laenen, F. Vander (2013). Gebruik van testinstrumenten in psychiatrische deskundigenverslagen bij internering; een exploratief dossieronderzoek in het gerechtelijk arrondissement Gent. *Tijdschrift voor Psychiatrie, 55*(5), 337-347.

Coid, J., Kahtan, N., Gault, S., Cook, A., & Jarman, B. (2001). Medium secure forensic psychiatry services: Comparison of seven English health regions. *British Journal of Psychiatry, 178*(1), 55-61.

Coid, J., & Kathan, N. (2000). An instrument to measure the security needs of patients in medium security. *The Journal of Forensic Psychiatry, 11*(1), 119-134.

Coid, J., Yang, M., Ullrich, S., Zhang, T., Sizmur, S., Farrington, D., & Rogers, R. (2011). Most items in structured risk assessment instruments do not predict violence. *The Journal of Forensic Psychiatry & Psychology, 22*(1), 3-21. doi: 10.1080/14789949.2010.495990

Cornaggia, C. M., Beghi, M., Pavone, F., & Barale, F. (2011). Aggression in psychiatry wards: A systematic review. *Psychiatry Research, 189*(1), 10-20. doi: 10.1016/j.psychres.2010.12.024

Crocker, A. G., Nicholls, T. L., Seto, M. C., & Côté, G. (2015). The national trajectory project of individuals found not criminally responsible on account of mental disorder in Canada. *Canadian Journal of Psychiatry, 60*(3), 96-97. doi: 10.1177/070674371506000303

Cullen, A. E., Jewell, A., Tully, J., Coghlan, S., Dean, K., & Fahy, T. (2015). A prospective cohort study of absconsion incidents in forensic psychiatric settings: Can we identify those at high-risk? *PLoS One, 10*(9), e0138819. doi: 10.1371/journal.pone.0138819

Daffern, M., Duggan, C., Huband, N., & Thomas, S. (2008). The impact of interpersonal style on aggression and treatment non-completion in patients with personality disorder admitted to a medium secure psychiatric unit. *Psychology, Crime & Law, 14*(6), 481-492. doi: 10.1080/10683160801948717

Decaire, M. W., Bedard, M., Riendeau, J., & Forrest, R. (2006). Incidents in a psychiatric forensic setting: Association with patient and staff characteristics. *Canadian Journal of Nursing Research, 38*(3), 68-80.

Delva, J. (1999). *De commissie internering. Eindverslag van de werkzaamheden*. Unpublished government report, Ministry of Justice.

DeMatteo, D., Edens, J. F., Galloway, M., Cox, J., Smith, S. T., Koller, J. P., & Bersoff, B. (2014). Investigating the role of the Psychopathy Checklist–Revised in United States case law. *Psychology, Public Policy, and Law, 20*(1), 96-107. doi: 10.1037/a0035452

Dinwiddie, S. H., & Briska, W. (2004). Prosecution of violent psychiatric inpatients: Theoretical and practical issues. *International Journal of Law and Psychiatry, 27*(1), 17-29. doi: 10.1016/j.ijlp.2004.01.002

Dirks-Linhorst, P. A., & Linhorst, D. (2006). A description of the design and costs of an insanity acquittee conditional release, monitoring, and revocation system. *International Journal of Forensic Mental Health, 5*(1), 55-65. doi: 10.1080/14999013.2006.10471230

Dolan, M., & Khawaja, A. (2004). The HCR–20 and post-discharge outcome in male patients discharged from medium security in the UK. *Aggressive Behavior, 30*(6), 469-483. doi: 10.1002/ab.20044

Edens, J. F., Colwell, L. H., Desforges, D. M., & Fernandez, K. (2005). The impact of mental health evidence on support for capital punishment: Are defendants labeled psychopathic considered more deserving of death? *Behavioral Sciences & the Law, 23*(5), 603-625. doi: 10.1002/bsl.660

Edens, J. F., Cox, J., Smith, S. T., DeMatteo, D., & Sorman, K. (2015). How reliable are Psychopathy Checklist-Revised scores in Canadian criminal trials? A case law review. *Psychological Assessment, 27*(2), 447-456. doi: 10.1037/pas0000048

Edens, J. F., Davis, K. M., Fernandez Smith, K., & Guy, L. S. (2013). No sympathy for the devil: Attributing psychopathic traits to capital murderers also predicts support for executing them. *Personality Disorders: Theory, Research, and Treatment, 4*(2), 175-181. doi: 10.1037/a0026442

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: Systematic review and meta-analysis. *The British Medical Journal, 345*. doi: 10.1136/bmj.e4692

Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. New York, NY: John Wiley.

Freestone, M., Taylor, C., Milsom, S., Mikton, C., Ullrich, S., Phillips, O., & Coid, J. (2012). Assessments and admissions during the first 6 years of a UK medium secure DSPD service. *Criminal Behaviour and Mental Health, 22*(2), 91-107. doi: 10.1002/cbm.1823

Gannon, T. A., & Ward, T. (2014). Where has all the psychology gone? A critical review of evidence-based psychological practice in correctional settings. *Aggression and Violent Behavior, 19*, 435-446. doi: 10.1016/j.avb.2014.06.006

Goethals, J. (1985). Internering in België. Enkele statische gegevens. *Panopticon, 6*(2), 129-147.

Gow, R. L., Choo, M., Darjee, R., Gould, S., & Steele, J. (2010). A demographic study of the Orchard Clinic: Scotland's first medium secure unit. *The Journal of Forensic Psychiatry & Psychology, 21*(1), 139-155. doi: 10.1080/14789940903188956

Gowensmith, W. N., Bryant, A. E., & Vitacco, M. J. (2014). Decision-making in post-acquittal hospital release: How do forensic evaluators make their decisions? *Behavioral Sciences & the Law, 32*(5), 596-607. doi: 10.1002/bsl.2135

Gradillas, V., Williams, A., Walsh, E., & Fahy, T. (2007). Do forensic psychiatric inpatient units pose a risk to local communities? *The Journal of Forensic Psychiatry & Psychology, 18*(2), 261-265. doi: 10.1080/14789940701246360

Gray, N. S., Snowden, R. J., MacCulloch, S., Phillips, H., Taylor, J., & MacCulloch, M. J. (2004). Relative efficacy of criminological, clinical, and personality measures of future risk of offending in mentally disordered offenders: A comparative study of HCR-20, PCL:SV, and OGRS. *Journal of Consulting and Clinical Psychology, 72*(3), 523-530. doi: 10.1037/0022-006X.72.3.523

Gray, N. S., Taylor, J., & Snowden, R. J. (2008). Predicting violent reconvictions using the HCR-20. *The British Journal of Psychiatry, 192*(5), 384-387. doi: 10.1192/bjp.bp.107.044065

Gudjonsson, G., Rabe-Hesketh, S., & Wilson, C. (2000). Violent incidents on a medium secure unit: The target of assault and the management of incidents. *The Journal of Forensic Psychiatry, 11*(1), 105-118. doi: 10.1080/095851800362391

Habets, P., Delespaul, P., & Jeandarme, I. (2016, September). *Experience sampling in de forensische psychiatrie: Meten van ervaringen in het dagelijks leven*. Paper presented at the Poster presented at the Eight Vlaams Geestelijk Gezondheidscongres, Antwerp, Belgium.

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist*. Toronto, ON: Multi-Health Systems.

Harte, J. M., Leeuwen, M. E. van, & Theuws, R. (2013). Agressie en geweld tegen hulpverleners in de psychiatrie; aard, omvang en strafrechtelijke reactie. *Tijdschrift voor Psychiatrie, 55*(5), 325-335.

Hayes, H., Kemp, R. I., Large, M. M., & Nielssen, O. B. (2014). A 21-year retrospective outcome study of New South Wales forensic patients granted conditional and unconditional release. *The Australian and New Zealand Journal of Psychiatry, 48*(3), 259-282. doi: 10.1177/0004867413507610

Heilbrun, L., & Griffin, P. A. (1993). Community-based forensic treatment of insanity acquittees. *International Journaf of Law and Psychiatry, 16*, 133-150.

Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology, 3*, 139–170. doi: 10.1111/j.2044-8333.1998.tb00355.x

Hildebrand, M., & Ruiter, C. de (2004). PCL-R psychopathy and its relation to DSM-IV Axis I and II disorders in a sample of male forensic psychiatric patients in The Netherlands. *International Journal of Law and Psychiatry, 27*(3), 233-248. doi: 10.1016/j.ijlp.2004.03.005

Jeandarme, I. (2016). Forensische psychiatrie à la flamande. *Orde van de Dag, 74*, 32-40.

Kumar, S., Fischer, J., Ng, B., Clarke, S., & Robinson, E. (2006). Prosecuting psychiatric patients who assault staff: A New Zealand perspective. *Australasian Psychiatry, 14*(3), 251-255. doi: 10.1111/j.1440-1665.2006.02296.x

Leistico, A. M., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the hare measures of psychopathy to antisocial conduct. *Law and Human Behavior, 32*(1), 28-45. doi: 10.1007/s10979-007-9096-6

Linhorst, D. M., & Scott, L. P. (2004). Assaultive behavior in state psychiatric hospitals: Differences between forensic and nonforensic patients. *Journal of Interpersonal Violence, 19*(8), 857-874. doi: 10.1177/0886260504266883

Lloyd, C. D., Clark, H. J., & Forth, A. E. (2010). Psychopathy, expert testimony, and indeterminate sentences: Exploring the relationship between Psychopathy Checklist-Revised testimony and trial outcome in Canada. *Legal and Criminological Psychology, 15*(2), 323-339. doi: 10.1348/135532509X468432

McMurran, M., & Theodosi, E. (2007). Is treatment non-completion associated with increased reconviction over no treatment? *Psychology, Crime, & Law, 13*(4), 333-343. doi: 10.1080/10683160601060374

Melzer, D., Tom, B., Brugha, T., Fryers, T., Gatward, R., Grounds, A., . . . Meltzer, H. (2004). Access to medium secure psychiatric care in England and Wales. 1: A national survey of admission assessments. *The Journal of Forensic Psychiatry & Psychology, 15*(1), 7-31. doi: 10.1080/14789940410001661856

Menghini, M., Ducro, C., & Pham, T. H. (2005). Évaluation de la récidive auprès d'agresseurs sexuels issus d'un hôpital sécuritaire en Belgique francophone. *Psychiatrie et Violence, 51*(1), 39-48.

Miller, C. S., Kimonis, E. R., Otto, R. K., Kline, S. M., & Wasserman, A. L. (2012). Reliability of risk assessment measures used in sexually violent predator proceedings. *Psychological Assessment, 24*(4), 944-953. doi: 10.1037/a0028411

Mudde, N., Nijman, H., Hulst, W. van der, & Bout, J. van den (2011). Het voorspellen van agressie tijdens de behandeling van forensisch psychiatrische patiënten aan de hand van de HCR-20 [Predicting aggression during the treatment of forensic psychiatric patients by means of the HCR-20]. *Tijdschrift voor Psychiatrie, 53*(10), 705-713.

Murrie, D. C., Boccaccini, M. T., Caperton, J., & Rufino, K. (2012). Field validity of the Psychopathy Checklist-Revised in sex offender risk assessment. *Psychological Assessment, 24*(2), 524-529. doi: 10.1037/a0026015

Murrie, D. C., Boccaccini, M. T., Johnson, J. T., & Janke, C. (2008). Does interrater (dis)agreement on Psychopathy Checklist scores in sexually violent predator trials suggest partisan allegiance in forensic evaluations? *Law and Human Behavior, 32*(4), 352-362. doi: 10.1007/s10979-007-9097-5

Murrie, D. C., Boccaccini, M. T., Turner, D. B., Meeks, M., Woods, C., & Tussey, C. (2009). Rater (dis)agreement on risk assessment measures in sexually violent predator proceedings: Evidence of adversarial allegiance in forensic evaluation? *Psychology, Public Policy and Law, 15*(1), 19-53.

Neal, T., Miller, S. L., & Shealy, R. C. (2015). A field study of a comprehensive violence risk assessment battery. *Criminal Justice and Behavior, 42*(9), 952-968. doi: 10.1177/0093854815572252

Nijman, H. L. I., Muris, P., Merckelbach, H. L. G. J., Palmstierna, T., Wistedt, B., Vos, A. M., . . . Allertz, W. (1999). The Staff Observation Aggression Scale-Revised (SOAS-R). *Aggressive Behavior, 25*(3), 197–209.

Olver, M. E., & Wong, S. C. (2009). Therapeutic responses of psychopathic sexual offenders: Treatment attrition, therapeutic change, and long-term recidivism. *Journal of Consulting and Clinical Psychology, 77*(2), 328-336. doi: 10.1037/a0015001

Palmstierna, T., & Wistedt, B. (1987). Staff observation aggression scale, SOAS: Presentation and evaluation. *Acta Psychiatrica Scandinavica, 76*(6), 657-663.

Pedersen, L., Ramussen, K., & Elsass, P. (2012). HCR-20 violence risk assessments as a guide for treating and managing violence risk in a forensic psychiatric setting. *Psychology, Crime & Law, 18*(8), 733-743. doi: 10.1080/1068316X.2010.548814

Pham, T. H., Ducro, C., Marghem, B., & Réveillère, C. (2005). Évaluation du risque de récidive au sein d'une population de délinquants incarcérés ou internés en Belgique francophone [Prediction of recidivism among prison inmates and forensic patients in Belgium]. *Annales Médico-Psychologiques, 163*(10), 878-881. doi: 10.1016/j.amp.2005.09.013

Quanbeck, C. (2006). Forensic psychiatric aspects of inpatient violence. *The Psychiatric Clinics of North America, 29*(3), 743-760. doi: 10.1016/j.psc.2006.04.011

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Research voor Beleid, Maastrichts Europees Instituut voor Transnationaal Rechtswetenschappelijk Onderzoek. (1995). *Voorzieningen voor de behandeling en bejegening van psychisch gestoorde delinquenten; een internationaal vergelijkende studie*. Leiden: Research voor Beleid, METRO.

Rotter, M., Carr, W. A., Magyar, M., & Rosenfeld, B. (2011). From incarceration to community care: Structured assessment of correctional adaptation. *Journal of the American Academy of Psychiatry and the Law, 39*(1), 72-77.

Ruyver, B. De, & Goethals, J. (1991). Actiedag internering, 5 oktober 1991. *Panopticon*, 184-189.

Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the Psychopathy Checklist and Psychopathy Checklist—Revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice, 3*(3), 203–215. doi: 10.1111/j.1468-2850.1996.tb00071.x

Salize, H. J., & Dressing, H. (2007). Admission of mentally disordered offenders to specialized forensic care in fifteen European Union member states. *Social Psychiatry and Psychiatric Epidemiology, 42*(4), 336-342. doi: 10.1007/s00127-007-0159-2

Singh, J. P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences & the Law, 31*(1), 8-22. doi: 10.1002/bsl.2052

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499-513. doi: 10.1016/j.cpr.2010.11.009

Skeem, J. L., Manchak, S., & Peterson, J. K. (2011). Correctional policy for offenders with mental illness: Creating a new paradigm for recidivism reduction. *Law and Human Behavior, 35*(2), 110-126. doi: 10.1007/s10979-010-9223-7

Stassen, W., Habets, P., Mertens, A., Laender, J. De, & Jeandarme, I. (2014). The InReach project: From penitentiary to forensic hospital. *Therapeutic Communities: The International Journal of Therapeutic Communities 35*(3), 119-126. doi: 10.1108/TC-01-2014-0002

Sturup, J., Edens, J. F., Sorman, K., Karlberg, D., Fredriksson, B., & Kristiansson, M. (2014). Field reliability of the Psychopathy Checklist-Revised among life sentenced prisoners in Sweden. *Law and Human Behavior, 38*(4), 315-324. doi: 10.1037/lhb0000063

Sundby, S. E. (1997). Capital jury and absolution: The intersection of trial strategy remorse and the death penalty. *Cornell Law Review, 83*(6), 1557-1598.

To, W. T. (2015). *Assessment and treatment perspectives on offenders with disabilities* [Unpublished doctoral dissertation]. Ghent University, Ghent, Belgium.

Troquete, N. A., Brink, R. H. van den, Beintema, H., Mulder, T., Os, T. W. van, Schoevers, R. A., & Wiersma, D. (2013). Risk assessment and shared care planning in out-patient forensic psychiatry: Cluster randomised controlled trial. *The British Journal of Psychiatry, 202*(5), 365-371. doi: 10.1192/bjp.bp.112.113043

Vandenbroucke, M. (1981). Bemerkingen bij het functioneren van de internering. *Panopticon, 1*, 80-85.

Vandevelde, S., Soyez, V., Beken, T. Vander, Smet, S. De, Boers, A., & Broekaert, E. (2011). Mentally ill offenders in prison: The Belgian case. *International Journal of Law and Psychiatry, 34*(1), 71-78. doi: 10.1016/j.ijlp.2010.11.011

Veeken, F. C. van der, Bogaerts, S., & Lucieer, J. (2015). Patient profiles in Dutch forensic psychiatry based on risk factors, pathology, and offense. *International Journal of Offender Therapy and Comparative Criminology*, Advance online publication. doi: 10.1177/0306624X15619636

Vertommen, H., & Maesschalck, C. (2007). *Definitie en operationalisering van het concept "medium risk"*. Research report.

Vicenzutto, A., & Pham, T. H. (2015). Les condamnés qu'on interne: plus fous? plus malades. *Annales Médico-Psychologiques*(173), 629-630.

Vitacco, M. J., Vauter, R., Erickson, S. K., & Ragatz, L. (2014). Evaluating conditional release in not guilty by reason of insanity acquittees: A prospective follow-up study in Virginia. *Law and Human Behavior, 38*(4), 346-356. doi: 10.1037/lhb0000071

Vogel, V. de, & Ruiter, C. de (2006). Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime & Law, 12*(3), 321-336. doi: 10.1080/10683160600569029

Vojt, G., Thomson, L. D. G., & Marshall, L. A. (2013). The predictive validity of the HCR-20 following clinical implementation: Does it work in practice? *The Journal of Forensic Psychiatry & Psychology, 24*(3), 371-385. doi: 10.1080/14789949.2013.800894

Walters, G. D. (2003). Predicting institutional adjustment and recidivism with the Psychopathy Checklist factor scores: A meta-analysis. *Law and Human Behavior, 27*(5), 541-558.

Walters, G. D., Kroner, D. G., & DeMatteo, D. (2014). The impact of base rate utilization and clinical experience on the accuracy of judgments made with the HCR-20. *Journal of Forensic Psychology Practice, 14*(4), 288-301. doi: 10.1080/15228932.2014.941726

Wand, T. (2012). Investigating the evidence for the effectiveness of risk assessment in mental health care. *Issues in Mental Health Nursing, 33*(1), 2-7. doi: 10.3109/01612840.2011.616984

Wayland, K., & O'Brien, S. (2013). Deconstructing antisocial personality disorder and psychopathy: A guideline-based approach to prejudicial psychiatric labels. *Hofstra Law Review, 42*(2), 519-588.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence* (Version 2). Burnaby, BC: Simon Fraser University.

Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Desmarais, S. (2009). *Manual for the Short-Term Assessment of Risk and Treatability (START)*. Port Coquitlam, BC: Forensic Psychiatric Services Commission and St. Joseph's Healthcare.

Wiederanders, M. R. (1992). Recidivism of disordered offenders who were conditionally vs unconditionally released. *Behavioral Sciences & the Law, 10*, 141-148.

Wiederanders, M. R., Bromley, D. L., & Choate, P. A. (1997). Forensic conditional release programs and outcomes in three states. *International Journal of Law and Psychiatry, 20*(2), 249-257.

Yang, M., Wong, S. C., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin, 136*(5), 740–767. doi: 10.1037/a0020473

This dissertation consists of ten studies regarding internees referred to a medium security unit for treatment. It covers research on the first 10 years after the establishment of these three units in Flanders and comprises practically the entire Flemish interned population treated in medium security during that period.

PART I. LEGISLATIVE BACKGROUND

The dissertation started with a discussion on the legislation regarding offenders found not guilty by reason of insanity (NGRI) and the functioning of the review boards (Commissions for the Protection of Society or CPSs), who are responsible for the execution of the internment measure. This legal background information and the practical implications are important since medium security units are obliged to work within the ambit of the applicable law and (release) decision processes of the CPSs. Furthermore, it helps explain the ways in which internment differs from legal provisions in other countries, including the Dutch Entrustment Act.

In **Chapter 1**, the conditions for internment were analyzed. The perpetrator must have committed an offense described in statutory law as a felony or misdemeanor. He or she must be found in a state of insanity, serious mental disorder, or deficiency that renders him/her unable to control his/her actions at the time of the judicial decision. Furthermore, the offender must be considered dangerous to society. Internment is an indefinite protective measure and not a punishment, and it has a dual purpose, namely the protection of society *and* the medical-psychiatric treatment of the offender to reduce the risk of future offending. The act was quite innovative when it was passed in 1930. Up until the beginning of the 20th century, offenders with mental illnesses were either punished or, when entirely deprived of their senses, acquitted on the basis of the insanity clause. However, several problems arose. First, criminal justice practice and the social defense literature held that many offenders could neither be regarded as fully responsible nor as criminally insane. Second, punishing them did not lead to behavioral changes, and when released they continued to be a threat to society. The need to adopt a new approach, i.e., a preventive measure to protect society following insanity acquittal, led to the passage of the Internment Act. As was the case with the law on conditional release (*Wet Lejeune*), the Internment Act was heavily influenced by the so called modern movement of criminal justice theories.

Internationally, the act was assessed as a good law, whereas criticisms were primarily related to the execution phase. There was an astonishing lack of adequate treatment providers, leaving a substantial number of internees untreated within the prison system. It took many years of continued pressure from people working in the field and courage from politicians to finally start the establishment of a categorical forensic care system.

Chapter 2 focused more specifically on the psychiatric expert reports and criminal responsibility assessments. The Internment Act however did not, and still does not, regulate forensic psychiatric expert advice, apart from the general provision that an expert witness can advise the court as a technical consultant. The advice is neither necessary nor mandatory for imposing an internment measure. However, consistent with the case law of the European Court of Human Right on the (un)lawful deprivation of liberty of a person unsound of mind, judges generally appoint psychiatric experts whenever internment decisions are at stake. The goal of the psychiatric observation is to evaluate the mental state of the accused. In practice, psychiatric reports have been heavily criticized, often because of highly contested cases in which conflicting psychiatric assessments are presented. Therefore, in this chapter, a systematic literature search was performed to assess the differences between a number of countries (the Netherlands, Canada, France, Sweden, Germany, and England) with respect to legal frameworks and procedures for conducting responsibility assessments. Belgium seems to be the only country that does not provide for the possibility of clinical observation. While there is a legal provision for establishing a specifically designed clinical observation center, this was never implemented. Belgium is also the only country that does not require a psychiatric assessment when mandatory treatment will be imposed. Criteria for both experts and expert reports are lacking, whereas examples can be found in other countries. A judicial reform on the internment measure is currently in progress and there is ongoing debate about the necessary criteria for both experts and expert reports.

PART II. MEDIUM SECURITY POPULATION

The second part of the dissertation provided a detailed overview of the entire study population that was treated in one of the medium security units (Chapter 3). In addition, Chapter 4 focused more specifically on the population that was administratively interned.

Chapter 3 began with a discussion on the risk and security level attributed to the medium security population. Risk, care, cure, and security are labels that are interrelated and sometimes interchangeably used. It is therefore important to be clear on what is meant when discussing medium security units. Descriptive analyses of the study population (*N* = 531) showed that the prototypical medium security patient is a male with a long criminal and psychiatric history prior to receiving an internment measure for a violent offense. Prior to first admission, he has been detained in prison one or several times for an average period of almost five years. He has poor social networks, lives alone and has no partner or children at the time of the index offense. He is poorly educated and has an unstable work history. He is diagnosed with a combination of psychosis, substance misuse, personality disorder and borderline intellectual functioning. Medium security treatment lasted on average 1.9 years. After medium security treatment, about 40% received continued treatment in a lower security unit. One-third of the patients dropped out prematurely from treatment, and more had their treatment course interrupted because of one or more time-out periods. Notwithstanding the difficulties inherent to international comparisons (i.e., due to differing legal systems and the related organization of forensic health care), a number of similar characteristics were found (e.g., with respect to demographic variables, psychiatric histories, index offenses, and levels of comorbidity). Differences were found with respect to clinical diagnoses, with a higher number of personality disorders and substance misuse found in the Flemish population, as well as a lower number of psychotic disorders. Also, the judicial histories of the internees were more pronounced and the risk assessment instruments scores were higher for the Flemish population. These differences may explain why discharge to the community was less frequent in comparison to other studies, and readmissions to medium security were higher than or comparable with other research. In terms of treatment characteristics, there were important drop-out rates. It could be hypothesized that the

population at hand was not always adequately matched to the offered medium security level, with at least a subpopulation needing a higher security level.

Chapter 4 discussed the controversial *Art. 21 procedure* of the internment law. This article holds that, in addition to regular internment measures, the Minister of Justice can impose an internment measure on a convict when, during detention, the convict is found to be "in a state of insanity, a serious state of mental disturbance or mental deficiency, which renders him/her unfit to control his/her actions". The procedure takes place without an adversarial debate and has been criticized for the weak legal position of the convicted internee; it has been abolished in the newly proposed law. The chapter then examined the implications of this new proposal for the regular mental health care system, thereby touching on a broader debate that deals with the connection between regular mental health care and forensic care. It was examined whether the clinical and risk profile of the population of convicted internees was compatible with non-categorical care. Chi-square tests and Fisher Exact tests were used to compare the group of convicted internees (n = 48) with the group of regular internees (n = 483) in case of categorical variables. In cases of continuous variables, t-tests or Mann-Whitney U tests were used. Compared to regular internees, convicted internees had similar psychopathology but, higher rates of psychoses not otherwise defined, which could have been prison acquired psychoses. Also, a higher risk for recidivism was found in the convicted internees. This became evident when analyzing the judicial histories (younger age at first conviction, more convictions, as well as more and longer detention periods) and the risk profiles (which were based on risk assessment instruments). Generally, these findings were in line with previous research on convicted internees in Wallonia. Further research is needed to determine whether the present results generalize to the entire group of convicted internees, for whom, as yet, no prevalence rates or profiles are available. As for now, abolishing the Art. 21 procedure would pose a serious challenge to regular psychiatric hospitals, especially in an era of de-institutionalization and cost cutting. Regular psychiatric units currently do not have the required – infrastructural and relational – security levels, nor do they have the skills to deal with this population.

PART III. INPATIENT MEDIUM SECURITY TREATMENT

Forensic psychiatric patients – and more specifically patients with psychopathic features – have traditionally been stigmatized as more violent, more difficult to treat and less compliant than other patients. As a result, general psychiatric institutions are reluctant to treat these patients and local communities are opposed to the presence of forensic units because they are concerned for the public safety (e.g., after absconding). The three inpatient studies aimed to examine if these presumptions are correct. The third part of the dissertation therefore examined which incidents were reported to the CPS during inpatient medium security treatment (Chapter 5), to identify the risk factors for inpatient violence (Chapter 6), and which incidents occurred in patients with psychopathic traits (Chapter 7). Both incidents occurring in the medium security unit and during (supervised) leave from the unit were analyzed.

Chapter 5 examined all incidents and the subsequent judicial reactions to these incidents by examining descriptive statistical analyses in the entire study population (*N* = 531). Chi-square tests were used to compare revocation rates between incident reports for violent and non-violent incidents. The study further added to the scant research on non-violent incidents and the quasi non-existent research on judicial reactions to incidents within a forensic unit. Despite the negative public perception of forensic patients that general psychiatric facilities and local communities hold, nearly half of the population had no incident reported to the CPS. The most frequently registered incidents were non-violent incidents such as absconding and treatment non-compliance. Internees were often granted (un)supervised leave quite soon after admission for a variety of reasons, such as attending treatment sessions that were held on the premises. As was shown, they also frequently absconded, albeit for a relatively short period (two-thirds returned within two days) and with offense-related behavior on only one out of five occasions. Drug-related incidents were found in more than half of the population. Together with the high percentage of substance misuse diagnoses, this finding highlighted the importance of focusing on ongoing substance misuse treatment or access to addiction services. The base rate for physically violent incidents was low. The study confirmed that the risk posed to local communities was small and the low institutional physical violence rate suggested adequate risk management by the staff. Crime-related incidents during medium security treatment were rarely

prosecuted and adjudicated. However, the base rate of revocation – and hence drop-out from treatment – as a result of incidents was high. Two-thirds of the patients who involved in an incident were incarcerated at least once due to revocation. The imprisonment periods following revocation were lengthy, lasting on average 1.8 years. Further research is needed to determine the reasons for these high drop-out rates. It can be assumed that the security level of the units was not adequately matched to the risk level posed by the patients and/or treatment was not responsive to characteristics of the patients (e.g., low intelligence levels). Continuing investment in therapeutic relations and other strategies to promote adherence is recommended. Also, maintaining careful records of incidents can increase awareness of risk factors and situations, inform policy decisions regarding aggression management on the facility level and eventually prevent more serious incidents in the future. At this moment a clear policy regarding the reporting of incidents – particularly those of an aggressive nature – during the treatment of medium security forensic psychiatric patients in Flanders is lacking. Therefore, a clear, consistent policy regarding the prosecution of patients is recommended.

Chapter 6 focused more specifically on verbally and physically violent incidents. Besides affecting staff, patients perpetrating violence are adversely affected as well, since seclusion and restraints are often used to manage aggressive behavior and previous research found that inpatient violence can result in prolonged involuntary confinement. The scant research on inpatient violence in the literature is surprising, because an important objective of forensic psychiatric treatment is to teach violent patients how to de-escalate. Furthermore, inpatient violence predicts violent recidivism after treatment. Besides the demographic, clinical, and criminogenic individual risk factors already identified in prior research, the study aimed to analyze whether interpersonal violence (IPV) would be associated with treatment drop-out. To that end, subsamples with several severity levels of interpersonal violence (i.e., none, verbal, physical, and a combination of both verbal and physical) were compared (N = 531), using the Chi-square or Fisher Exact test in case of categorical variables and the one-way analysis of variance (ANOVA) for continuous variables. Next, significant bivariate associations were entered in multiple logistic regression analyses. Results showed that interpersonally violent patients misbehaved more generally during treatment (i.e., more violent patients absconded, did not comply with hospital rules, and kept drinking alcohol during treatment even though this act was

prohibited). In fact, after controlling for other variables, treatment characteristics were the only characteristics that predicted interpersonal violence, whereas the commonly used risk assessment instruments and other well-established risk factors for violence, such as previous violence, were not found to be useful predictors. There was also an increased risk for drop-out, both in the physically violent patients (67%), as in the verbally violent patients (60%). It therefore remains important to be vigilant for verbal violence and investigate this form of violence further. Taken together, these results suggest that dynamic treatment variables can act as warning signals for IPV. Furthermore, it is again noted that monitoring treatment processes and inpatient incidents preferably with the use of standardized instruments, is important. Also, the need for a more dynamic and contextual approach to investigating predictors of inpatient IPV was highlighted by the few individual risk factors that showed different associations with verbal and physical IPV in the present study. Indeed, triggers for violence are multifactorial and involve an interplay between individual, situational and structural factors.

Chapter 7 described the treatment course of a subpopulation of internees, namely those with high psychopathic traits. According to the Risk Need Responsivity model, intensive care and supervision is required for patients with high psychopathic traits. However, there is much reluctance to take these patients into treatment, not only because of the perceived limited chances of success but also because therapy-interfering behavior is expected. When the first medium security units in Flanders were established, primary psychopathy was one of the exclusion criteria. Notwithstanding this exclusion criterion, internees with comorbid medium and high psychopathy traits were accepted for treatment. In this chapter, internees that were assessed with the Psychopathy Checklist-Revised (PCL-R; *N* = 224) were divided into three groups: low, medium, and high psychopathy traits. Associations between psychopathy and criminogenic risk and need factors were analyzed. In addition, the association between psychopathy and therapy-interfering behavior (defined as non-compliance, drop-out, and institutional misconduct) was investigated with correlational and logistic regression analyses. Results showed that psychopathy was associated with greater risk and needs and therapy-interfering behavior. PCL-R Factor 2 predicted institutional misconduct, whereas PCL-R Factor 1 predicted drop-out from treatment. Taken together, this study thus confirmed other research stating that treatment should focus on criminogenic PCL-R Factor 2 features but must also carefully take into account

PCL-R Factor 1 characteristics in order to avoid drop-out. Again, the study highlighted the importance of a responsive treatment climate to keep this difficult-to-treat group in treatment.

PART IV. RISK ASSESSMENT DURING AND AFTER MEDIUM SECURITY TREATMENT

Part IV of the dissertation covered an important topic in forensic psychiatry, namely the assessment of the risk of violent reoffending. Structured risk assessment has become part of routine practice in forensic settings. However, little attention has been paid to the clinical applicability of existing risk assessment tools and their field validity.

In **Chapter 8,** the predictive validity of one of the most commonly used risk assessment instruments, the Historical, Clinical, Risk management-20 (HCR-20) was studied. Predictive validity is usually measured using the Area Under the Curve (AUC) as an outcome measure, which indicates the probability that a randomly selected recidivist has a higher risk classification than a randomly selected non-recidivist. However, this retrospective measure does not provide a full picture of the predictive value of a risk assessment tool. It is therefore recommended to including both components of a tools' predictive validity, namely discrimination (i.e., AUC, sensitivity, and specificity) and calibration (i.e., positive predictive value, negative predictive value, number needed to detain, and number safely discharged). Prospectively-orientated statistics measuring how well a tool's prediction of risk agrees with the actual observed risk may provide additional useful information, in that they simulate clinical decision-making. HCR-20 scores coded within one year after medium security unit admission (n = 168) and within one year ahead of discharge (n = 105) from the unit in the period 2001–2010 were included in this study. Recidivism was evaluated twice: while the internee was still a medium security inpatient (both referring to incidents occurring on the hospital premises and during leave or when absconding), and after discharge from the medium security unit. Predictions of violent recidivism during and after medium security treatment using AUC were non-significant. In addition, analyses showed that the HCR-20 was mainly useful for identifying low-risk individuals. Possible explanations for the poor predictive validity were discussed (e.g., adequate risk management affects predictive accuracy). Ultimately, the main goal of risk assessment is not predicting future violence but rather reducing this risk.

Chapter 9 investigated the interrater reliability and predictive validity of the PCL-R when assessed by raters in the field. Internationally, there has been a small but accumulating body of literature that suggests considerably attenuated reliability and predictive validity when the PCL-R is used in applied forensic settings. Research focusing on rater differences in applied settings has focused primarily on two non-mutually exclusive explanations for these differences: adversarial allegiance and individual differences across raters. The potential adverse implications of these field validity findings should not be taken lightly, given concerns about the stigmatizing effects of the *psychopath* label on legal decision-making. However, the total number of field validity studies is relatively limited (particularly in relation to examinations of factor scores), and sometimes based on small samples in particular proceedings such as the Sexually Violent Predator trials in the US. As such, examining the reproducibility of these results, particularly in jurisdictions outside of North America and with samples other than sex offenders, is particularly important. The current study investigated scoring discrepancies between hospital and prison settings, as well as differences in predictive validity across these two settings. PCL-R information was collected from prison and hospital files, resulting in 224 PCL-R total scores and 74 double scores. Although we did not expect adversarial allegiance per se, we assumed that the study might show evidence of some kind of partisanship or contextual pressures that might impact scores across these settings. For example, given systemic pressures within the prison system, it could be argued that prison evaluators might feel more pressure to score an examinee lower on the PCL-R than the corresponding evaluators in forensic institutions, in order to facilitate the transfer of forensic patients from prison to hospital. When examining repeated measurements, large individual differences were found together with an intraclass correlation coefficients ($ICC_{A,1}$) of .42 for the total score. Discrepant results were found for PCL-R Factor 2, with repeated scores within the same setting having an $ICC_{A,1}$ of .28 versus an $ICC_{A,1}$ of .57 for repeated scores between settings. However, the AUCs for the PCL-R Total, factor, and facet scores did not differ between settings. For the whole sample, PCL-R Factor 2 scores marginally predicted (violent) recidivism after two years (AUC = .62 and .63), whereas PCL-R Factor 1 did not predict (violent) recidivism. Consistent with recent studies from other countries, these results suggest inadequate field reliability and validity in prison and hospital settings in Flanders. The clinical implications of these findings were discussed and included for example using only prolific raters to score the PCL-R.

PART V. RECIDIVISM AFTER MEDIUM SECURITY TREATMENT

Treatment outcome in forensic mental health settings is best measured over a broad range of areas, including clinical and humanitarian ones. However, preventing future criminal behavior is the most important goal in forensic psychiatric treatment. It is therefore surprising that no recidivism data existed on Flemish internees prior to the current research. Internationally, in adult forensic populations, recidivism rates varied from 7.1% to 63% for general recidivism and 1.8% to 46.0% for violent recidivism over different follow-up periods, and revocation rates of rule violations ranged from 5% to 49%.

In **Chapter 10,** recidivism data based on criminal records were collected from the Central Criminal Records of the Ministry of Justice. In order to examine the effect of conditional release on recidivism, reconviction data of two subgroups (internees under conditional release and internees who received unconditional release) were compared. Furthermore, in the unconditional release group, a comparison between the conditional release period and the unconditional release period was made. Recidivism rates were reassuringly low; 12.5% of the internees recidivated with a general offense and 7.4% with a violent offense, whereas none of the unconditionally released patients reoffended for a serious crime. Furthermore, there was a decrease between pre-treatment and post-treatment offending. Although these results are encouraging, without a RCT, it remains unclear whether these changes can be attributed to treatment or in comparison with a non-treatment condition (e.g., treatment as usual or waiting condition). When comparing the unconditional release group to the conditional group, and when comparing the conditional release periods to the unconditional release periods, it became clear that the reduction in danger level was most apparent when the interment measure was in effect (i.e., the patient was under some form of supervision). However, the difference was not statistically significant, probably reflecting low statistical power due to the small numbers. This study also examined revocation rates and revealed that post-treatment revocation rates were significantly higher than reconviction rates, which was in line with international research. These high revocation rates took place in protection-oriented conditional release programs, resulting in (longer) incarceration periods, and can be considered the downside of the low recidivism rates. Patients are sent back to prison without any formal trial, which further hampers their chances of

returning to a forensic unit (and, in a later stage, to general psychiatry). These re-incarcerations might further increase the negative stereotypes held by regular psychiatry and the distrust of forensic patients in (forensic) mental health services. Furthermore, because there is no prosecution and conviction, the rights of the potential victims are not taken into account by a criminal court. Future research is needed to determine whether revocations can be identified as a major cause of low re-offense rates. International research showed that the relationship between revocations and re-offenses was not linear. Overall, our study suggested that successful (conditional) release by CPS in Flanders has several key success features, such as centralized responsibility, a uniform system of supervision by judicial officers, flexible procedures, and a close collaboration with the forensic treatment facilities, resulting in safe release plans. Currently another new internment law is again under revision. Further research is needed to evaluate the (criminal) outcomes of this more legal rights-based model with additional bureaucratic procedures and due process.

Chapter 11 assessed recidivism in a subpopulation, namely medium security internees under the jurisdiction of CPS Ghent (*N* = 192). In Chapter 10, official recidivism data were chosen as the primary outcome measure for recidivism because these are considered to be the most reliable source. However, this measure results in underreporting because re-arrest rates and dark numbers are not taken into account. This may be even more problematic in internees, since it was expected that crime-related incidents would not automatically be prosecuted. Therefore, crime-related incidents (defined as incident coded under offending categories of the Belgian penal code, whether or not they led to further prosecution or sentencing) that were reported to the CPS were examined in order to provide a more detailed picture of recidivism. Significantly more internees had incidents reported than official recidivism, and the fivefold increase in both general and violent relapse was higher than expected. The lack of prosecution can be explained by the mandatory supervision of the CPS, which allows the prosecutor to re-incarcerate the internee within a rather flexible procedure without contradictory debate; moreover this procedure is immediately carried out, thereby avoiding the lengthy and lingering procedure of a new trial, which would in most cases result in yet another internment measure. The present study also revealed that most crime-related incidents were less serious in nature, which might be another reason to avoid prosecution. The study's secondary aim was to analyze risk factors for

recidivism. To that end, the group that recidivated (*n* = 43) was compared to the group that did not recidivate (*n* = 102) with a violent offense; the risk factors used for comparison were identified in the international literature. The association between relapses in a violent crime, and each of the personal, offense and treatment-related characteristics were individually examined by means of the Pearson chi-square and Fisher exact test for categorical data and with the unpaired t-test or Mann Whitney U tests for continuous data. Characteristics that had a significant association with violent crime relapse were then entered into multiple logistic regression. It was expected that historical factors such as age at first sentence or number of prior sentences would better predict violent recidivism than clinical factors such as diagnosis of a major psychiatric disorder. When the characteristics associated with violent recidivism were entered in one regression model, only the variable habitual offender was independently associated with recidivism in violent crime. The low base rate of recidivism – and subsequent loss of statistical power – could be an explanation for the limited links that were found. Further research with the entire population is therefore recommended.

PART VI. VICTIMS

This dissertation ended with a chapter on victims. Ultimately, this is the goal of forensic psychiatric treatment, namely the prevention of new victims. Knowledge on victim characteristics (e.g., age, gender, and relationship to the offender) enables more effective treatment of internees and the prevention of future victims. Surprisingly, in contrast to the scholarly attention paid to forensic perpetrators, studies examining victim characteristics are scarce.

In **Chapter 12,** victims of the index internment measure were described. Information on these offenses was retrieved from the Central Criminal Records of the Ministry of Justice and cross-referenced with the different court administrations. For each internee with a violent index offense, the following victim variables were analyzed: number of victims, gender of the victim, whether the victim was a minor or adult, and whether the victim was a stranger or acquaintance. Ten participants were removed because all information on victims was missing, leaving an analytical study population of *N* = 362. Repeated measures ANOVA analyses were used for parametric data and Wilcoxon signed rank tests or Friedman tests were used for non-parametric

data. The results were in line with previous research: adult victims who were known to the offender were most frequently victimized. In addition, family victims were the most frequent victims among the acquaintance victims. No gender differences emerged. It was further investigated whether this victim profile would differ regarding gender, age, and relationship of the victims after stratification for the psychiatric diagnoses of the offender (psychotic vs. personality disorder vs. psychotic and personality disorder vs. other)? It was found that these victim profile characteristics were highly similar for perpetrators with different psychiatric diagnoses. Together, these results may inform and help practitioners recognize potentially risky situations for caregivers and family members. Although violence committed by patients with a major mental disorder is in general a phenomenon with a low base rate, preventive actions can be taken on several levels, as the most likely victims will be acquaintances.

The major findings, limitations and strengths were addressed in a **general discussion**. A critical analysis of the thesis was presented, along with some recommendations for further research.

Dit proefschrift bestaat uit tien studies met betrekking tot de behandeling van geïnterneerden op een medium security unit. Het onderzoekt de eerste 10 jaar na de oprichting van de drie medium security eenheden in Vlaanderen, en bestrijkt vrijwel de gehele behandelde populatie.

DEEL I. WETGEVEND KADER

Het proefschrift begon met een uiteenzetting over de wetgeving inzake daders die niet schuldig bevonden werden wegens ontoerekeningsvatbaarheid, en over de werking van de Commissies ter Bescherming van de Maatschappij (CBM), verantwoordelijk voor de uitvoering van de interneringsmaatregel. Deze juridische achtergrondinformatie en haar praktische implicaties zijn belangrijk omdat medium security units verplicht zijn te werken binnen het toepasselijke recht en de besluitvormingsprocessen van de CBM. Bovendien helpt het wetgevend kader te begrijpen op welke aspecten het interneringssysteem afwijkt van wettelijke bepalingen in andere landen, zoals het systeem van Terbeschikkingstelling (TBS) in Nederland.

In **Hoofdstuk 1** werden de voorwaarden voor internering geanalyseerd. De dader moet een strafbaar feit hebben gepleegd, wettelijk omschreven als een wanbedrijf of misdaad. Hij of zij moet zich "hetzij in staat van krankzinnigheid, hetzij in een ernstige staat van geestesstoornis, of van zwakzinnigheid" bevinden die hem/haar ongeschikt maakt tot het controleren van zijn/haar daden op het moment van de rechterlijke beslissing. Bovendien moet de dader als maatschappelijk gevaarlijk worden beschouwd. Internering is een beschermingsmaatregel van onbepaalde duur en geen straf, en heeft een tweeledig doel: de bescherming van de samenleving én de medisch-psychiatrische behandeling van de dader. De interneringswet was in 1930 heel innovatief. Tot aan het begin van de 20ste eeuw, werden delinquenten met psychische aandoeningen ofwel gestraft of, indien ze "volledig beroofd waren van hun zinnen", vrijgesproken op basis van de waanzin-clausule. Deze wetgeving leverde in de praktijk veel problemen op. Ten eerste bleek, zowel vanuit de strafrechtspraktijk als vanuit de zogenaamde sociaal verweer-literatuur, dat veel daders niet als volledig verantwoordelijk noch als crimineel krankzinnig konden worden beschouwd. Ten tweede bleek louter bestraffing niet te leiden tot gedragsverandering. Na vrijstelling bleken delinquenten nog steeds gevaarlijk voor de maatschappij. De noodzaak tot een nieuwe aanpak, een preventieve maatregel dus die de

maatschappij zou beschermen tegen daders die vrijgesproken werden wegens krankzinnigheid, leidde na veel discussies tot de eerste interneringswet. Net zoals bij de wet op de voorwaardelijke invrijheidstelling (wet Lejeune), was de interneringswet (ook wel wet op sociaal verweer genoemd) sterk beïnvloed door de zogenaamde moderne beweging in de strafrechtstheorie. Internationaal werd de nieuwe wet positief geëvalueerd. De kritiek richtte zich voornamelijk op de uitvoeringsfase. Er was een verbazingwekkend gebrek aan adequate behandelsettings, waardoor een groot aantal geïnterneerden onbehandeld in de gevangenis bleef. Het duurde vele jaren van voortdurende druk van de mensen in het veld en politieke moed om uiteindelijk een categoraal zorgaanbod uit te bouwen.

Hoofdstuk 2 richtte zich meer specifiek op de psychiatrische rapporten van deskundigen en evaluaties met betrekking tot de strafrechtelijke verantwoordelijkheid. Wettelijk was en is dit advies nog steeds niet aan voorwaarden gebonden, afgezien van de algemene bepaling dat een deskundige de rechtbank kan adviseren als technisch adviseur. Het advies is noodzakelijk noch verplicht voor het opleggen van een interneringsmaatregel. In overeenstemming met de jurisprudentie van het Europees Hof voor de Rechten van de mens over de (on)rechtmatige vrijheidsbeneming van psychisch gestoorde personen, benoemen rechters over het algemeen echter toch psychiatrische experts wanneer zij een interneringsbeslissing overwegen. Het doel van de psychiatrische observatie is om de mentale toestand van de verdachte te evalueren. In de praktijk worden psychiatrische rapporten zwaar bekritiseerd, vaak na omstreden rechtszaken waarin tegenstrijdige psychiatrische beoordelingen voor het voetlicht worden gebracht. Daarom werd in dit hoofdstuk een systematisch literatuuronderzoek uitgevoerd, met de bedoeling de verschillen inzake wettelijke kaders en procedures voor het uitvoeren van de evaluaties omtrent toerekeningsvatbaarheid internationaal (Nederland, Canada, Frankrijk, Zweden, Duitsland en Engeland) te vergelijken. Goede praktijken uit het buitenland werden afgetoetst. België lijkt het enige land dat niet voorziet in de mogelijkheid van een klinische observatie. De wettelijke bepaling die voorziet in de oprichting van een speciaal ontworpen klinisch observatiecentrum, werd nooit in uitvoering gebracht. België is ook het enige land dat geen psychiatrisch onderzoek vereist vooraleer een verplichte behandeling op te leggen. Criteria voor zowel deskundigen als deskundigenrapporten ontbreken, terwijl er voorbeelden te vinden zijn in het buitenland. De

interneringswet is momenteel onderhevig aan een wetswijziging en de discussie over de noodzakelijke criteria voor zowel experts als deskundigenrapporten is in volle gang.

DEEL II. MEDIUM SECURITY POPULATIE

Het tweede deel van het proefschrift gaf in het derde hoofdstuk een gedetailleerd overzicht van de totale studiepopulatie die in één van de medium security units werd behandeld. Het vierde hoofdstuk richtte zich dan weer meer specifiek op de subpopulatie die administratief werd geïnterneerd.

Hoofdstuk 3 begon met een discussie over de risicoprofielen en beveiligingsniveaus gerelateerd aan het label medium security. Risico, behandeling, begeleiding en veiligheid zijn termen die nauw met elkaar verbonden zijn en soms door elkaar gebruikt worden. Daarom is het belangrijk duidelijk te omschrijven wat bedoeld wordt met medium security. Beschrijvende analyses van de onderzoekspopulatie (*N* = 531) toonden aan dat de prototypische medium security-patiënt een man is met een lange criminele en psychiatrische voorgeschiedenis vooraleer de interneringsmaatregel wordt opgelegd voor een geweldmisdrijf. Voorafgaand aan de eerste opname in een medium security unit, heeft hij één of meerdere detentieperiodes achter de rug met een gemiddeld gevangenisverblijf van bijna vijf jaar. Hij heeft een slecht sociaal netwerk, woont alleen en heeft geen partner of kinderen op het moment van het index delict. Hij is slecht opgeleid en heeft een instabiele werkgeschiedenis. Hij is gediagnosticeerd met een combinatie van psychose, drugmisbruik en persoonlijkheidsstoornis. Zijn intelligentie ligt op zwakbegaafd niveau. De medium security behandeling nam gemiddeld 1.9 jaar in beslag. Na medium security behandeling, ontving ongeveer 40% van de populatie voortgezette behandeling binnen een lager beveiligingsniveau. Een derde van de patiënten maakte de medium security behandeling niet af, bij nog een groter deel van de groep werd de behandeling onderbroken door time-outperiodes. Niettegenstaande de moeilijkheden inherent aan internationale vergelijkingen (als gevolg van verschillende juridische systemen en de daarmee verband houdende organisatie van de forensische zorg), werd een aantal vergelijkbare kenmerken gevonden in de internationale literatuur (bijvoorbeeld met betrekking tot demografische variabelen, psychiatrische geschiedenissen, index delicten, en comorbiditeit). Verschillen werden

vastgesteld inzake klinische diagnoses, met een groter aantal persoonlijkheidsstoornissen en middelenmisbruik in de Vlaamse medium securitypopulatie, en een kleiner aantal psychotische aandoeningen. Ook de justitiële antecedenten waren meer uitgesproken, de scores op de instrumenten voor risicotaxatie lagen gemiddeld hoger. Deze verschillen kunnen verklaren waarom bijvoorbeeld een ambulante reclassering minder vaak en heropnames even vaak of vaker voorkwamen in vergelijking met andere studies. Bij de behandelingskenmerken viel de belangrijke uitval of drop-out op. Een verklarende hypothese zou hier kunnen zijn dat tenminste een deel van de behandelde populatie idealiter een hoger beveiligingsniveau nodig had.

In **Hoofdstuk 4** werd de controversiële Art. 21 procedure van de interneringswet besproken. Dit artikel stelt dat, in aanvulling op de reguliere internering, de minister van Justitie een interneringsmaatregel kan opleggen aan een veroordeelde die zich tijdens zijn detentie blijkt te bevinden in "een staat van krankzinnigheid of in een ernstige staat van geestesstoornis of van zwakzinnigheid die hem/haar ongeschikt maakt tot het controleren van zijn/haar daden". De procedure vindt plaats zonder tegensprekelijk debat, en kreeg kritiek omdat de rechtspositie van de veroordeelde geïnterneerde te zwak zou zijn. De procedure werd afgeschaft in het voorstel tot nieuwe interneringswet. Het hoofdstuk onderzocht vervolgens de implicaties van dit nieuwe wetsvoorstel voor de reguliere geestelijke gezondheidszorg. Daarbij raakte het aan een breder debat over de raakvlakken en verbindingen tussen de reguliere geestelijke gezondheidszorg en forensische zorg. Er werd onderzocht of het klinisch profiel en het risicoprofiel van de veroordeelde geïnterneerden compatibel waren met niet-categorale zorg. Chi kwadraat-toetsen en Fisher Exact-testen werden gebruikt om de groep veroordeelde geïnterneerden (n = 48) met de groep reguliere geïnterneerden (n = 483) te vergelijken in geval van categorische variabelen. In geval van continue variabelen, werden t-testen of Mann-Whitney U-testen gebruikt. In vergelijking met reguliere geïnterneerden, bleken veroordeelde geïnterneerden een soortgelijke psychopathologie te hebben, maar werden wel vaker psychotische stoornissen teruggevonden, mogelijk te wijten aan effecten van de detentie. Ook werd er een hoger risico op herval vastgesteld bij de veroordeelde geïnterneerden. Dit werd duidelijk bij de analyse van de justitiële antecedenten (jongere leeftijd bij de eerste veroordeling, meer veroordelingen, maar ook meer en langere detentieperiodes) en de risicoprofielen (die gebaseerd waren op risicotaxatie-instrumenten). Over het algemeen waren de bevindingen in lijn met eerder onderzoek naar

veroordeelde geïnterneerden in Wallonië. Verder onderzoek is nodig om te bepalen of de huidige resultaten veralgemeend kunnen worden naar de gehele groep van de veroordeelde geïnterneerden, waarvoor prevalentiecijfers noch profielen beschikbaar zijn. Voorlopige conclusie is dat de afschaffing van de Art. 21 procedure een serieuze uitdaging zou betekenen voor reguliere psychiatrische ziekenhuizen, met name in een tijdperk van de-institutionalisering en kostenbesparingen. Reguliere psychiatrische eenheden beschikken momenteel niet over de vereiste – infrastructurele en relationele – beveiligingsniveaus, noch hebben zij de vaardigheden die hen in staat stellen om te gaan met deze populatie.

DEEL III. MEDIUM SECURITY OPNAMES

Forensisch psychiatrische patiënten – en meer in het bijzonder patiënten met psychopathische kenmerken – zijn van oudsher gestigmatiseerd als gewelddadiger, moeilijker te behandelen en minder geneigd zich te conformeren aan behandelingsafspraken dan andere patiënten. Hierdoor zijn algemene psychiatrische instellingen terughoudend om deze patiënten te behandelen, en zijn lokale gemeenschappen vaak gekant tegen de aanwezigheid van forensische eenheden omdat ze zich zorgen maken om de openbare veiligheid (bijvoorbeeld na een ontvluchting). De drie klinische studies in het derde deel van de thesis onderzochten of deze vermoedens gerechtvaardigd zijn. Hiervoor werd nader bekeken welke incidenten gemeld werden bij de CBM tijdens medium security behandeling (Hoofdstuk 5), welke risicofactoren geweldsincidenten voorspellen (Hoofdstuk 6), en of incidenten vaker voorkomen bij patiënten met psychopathische trekken (Hoofdstuk 7). Zowel incidenten op de medium security units zelf als incidenten tijdens (begeleid) verlof vanuit de units werden geanalyseerd.

Hoofdstuk 5 onderzocht aan de hand van beschrijvende statistische analyses alle incidenten en de justitiële reacties op deze incidenten in de totale onderzoekspopulatie (*N* = 531). Chi-kwadraat-testen werden gebruikt om na te gaan of er een verschil was tussen het intrekken van vrijstellingen op proef bij gewelddadige dan wel niet-gewelddadige incidenten. De studie wou verder een bijdrage leveren aan het schaarse onderzoek naar niet-gewelddadige incidenten en het quasi onbestaande onderzoek inzake justitiële reacties op incidenten binnen een forensische eenheid. Ondanks het negatieve stigma dat aan forensische patiënten wordt toegekend, zowel

door de publieke opinie als door de reguliere psychiatrie, verliep de behandeling bij bijna de helft van de populatie zonder dat incidenten gerapporteerd werden aan de CBM. De geregistreerde incidenten waren in de eerste plaats niet-gewelddadige incidenten, zoals onttrekkingen aan toezicht en non-compliance. Geïnterneerden krijgen vrij snel (on)begeleide vrijheden om een verscheidenheid aan redenen, zoals bijvoorbeeld het bijwonen van therapiesessies op het ziekenhuisterrein. Hoewel onttrekkingen vrij vaak voorkwamen, waren ze meestal van korte duur (twee derde keerde binnen twee dagen terug) en was er bij slechts één op de vijf onttrekkingen sprake van een bijkomend delict gerelateerd incident. Middelen gerelateerde incidenten werden vastgesteld bij meer dan de helft van de populatie. Samen met het hoge percentage diagnoses van middelenmisbruik, wees dit op het belang van behandelingen gericht op verslaving of het verzekeren van toegang tot verslavingszorg. De base rate voor fysiek gewelddadige incidenten was laag. De studie bevestigde dat het risico voor lokale gemeenschappen beperkt bleef en het risicomanagement van teams op de units adequaat was. Enerzijds werd vastgesteld dat incidenten die als misdrijf konden omschreven worden, zelden strafrechtelijk werden vervolgd en berecht. Anderzijds bleek er ten gevolge van incidentenrapportage vaak sprake van het herroepen van een vrijstelling op proef, en bijgevolg ook een drop-out van de behandeling. Twee derde van de patiënten verwikkeld in een incident, werden tenminste eenmaal opgesloten door herroeping van de vrijstelling op proef. De detentieperiodes die hierop volgden waren langdurig, gemiddeld 1,8 jaar. Verder onderzoek is nodig om de redenen voor deze hoge drop-outpercentages te bepalen. Voorlopige hypotheses gaan er van uit dat het beveiligingsniveau van de eenheden onvoldoende was afgestemd op het risiconiveau van de patiënt en/of dat de behandeling te weinig rekening hield met bepaalde kenmerken van de patiënten, zoals hun laag intellectueel niveau. Voortdurende investering in therapeutische relaties en andere strategieën om de compliance te bevorderen, worden aanbevolen. Ook het handhaven van een zorgvuldige registratie van incidenten kan bewustwording van risicofactoren en - situaties verhogen, en beleidsbeslissingen inzake agressiebeheer faciliteren, en uiteindelijk voorkomen dat meer ernstige incidenten in de toekomst plaatsvinden. Op dit moment ontbreekt in Vlaamse medium security units een duidelijk beleid als het gaat om het melden van incidenten, in het bijzonder die van een agressieve aard. Daarom is een duidelijk, consistent beleid ten aanzien van de vervolging van de patiënten aanbevolen.

Hoofdstuk 6 richtte zich meer in het bijzonder op verbaal en fysiek gewelddadige incidenten. Geweld heeft niet enkel een negatieve invloed op personeel, maar ook op de patiënten zelf, gezien isolaties en andere vrijheidsbeperkende maatregelen vaak getroffen worden tegenover gewelddadige patiënten en dit ook kan leiden tot langer durende gedwongen opname. Het schaars onderzoek naar institutioneel geweld in de literatuur is verrassend, omdat een belangrijke doelstelling van forensisch psychiatrische behandeling er net uit bestaat om escalaties bij gewelddadige patiënten te voorkomen. Bovendien voorspelt geweld tijdens opname ook recidive na behandeling. Naast onderzoek van individuele demografische, klinische en criminogene risicofactoren die in eerder onderzoek waren geïdentificeerd, had deze studie tot doel te analyseren of interpersoonlijk geweld (IPG) gepaard gaat met verhoogd risico op drop-out. Daartoe werden groepen met oplopende ernst van geweld (dat wil zeggen: geen, verbaal, fysiek, en een combinatie van zowel verbaal als fysiek geweld) vergeleken (*N* = 531), met behulp van de Chi-kwadraat of Fisher Exact-testen in geval van categorische variabelen en de one-way analyse van variantie (ANOVA) in geval van continue variabelen. Vervolgens werden significante bivariate associaties opgenomen in multipele logistische regressie-analyses. De resultaten toonden aan dat gewelddadige patiënten zich meer in het algemeen misdroegen tijdens de behandeling (meer gewelddadige patiënten onttrokken zich aan toezicht, waren non-compliant, en dronken alcohol tijdens de behandeling hoewel dit verboden was). Na correctie voor andere variabelen waren deze behandelingskenmerken de enige elementen die interpersoonlijk geweld voorspelden. En dat terwijl populaire risicotaxatie-instrumenten en andere gevestigde risicofactoren voor geweld, zoals voorafgaandelijk geweld, geen nuttige voorspellers bleken te zijn. Er was ook een verhoogd risico op drop-out, zowel bij fysiek gewelddadige patiënten (66.7%) als bij verbaal gewelddadige patiënten (60.0%). Het is daarom belangrijk ook waakzaam te blijven voor verbaal geweld en ook deze vorm van geweld verder te onderzoeken. In hun geheel suggereerden de resultaten dat dynamische behandelingskenmerken kunnen fungeren als waarschuwingssignalen voor IPG. Verder werd opgemerkt dat het monitoren van behandelingsvoortgang en incidenten tijdens de behandeling, bij voorkeur met gestandaardiseerde instrumenten, belangrijk is. Het feit dat in deze studie slechts in beperkte mate individuele voorspellers werden gevonden, benadrukte ook het belang van een meer dynamische en contextuele benadering bij onderzoek naar voorspellers van institutioneel IPG.

Triggers voor geweld zijn namelijk multifactorieel, en behelzen een wisselwerking tussen individuele, situationele en structurele factoren.

Hoofdstuk 7 beschreef het behandelingsverloop van een subpopulatie van geïnterneerden, namelijk diegenen met een hoge mate van psychopathie. Volgens het Risk Need Responsivity model, hebben patiënten met een hoge mate van psychopathie intensieve zorg en begeleiding nodig. Toch is er veel terughoudendheid om deze patiënten in behandeling te nemen, niet alleen vanwege de verwachte beperkte kansen op succes, maar ook omdat therapiestorend gedrag wordt verwacht. Toen de eerste medium security units in Vlaanderen werden opgericht, was kernpsychopathie één van de uitsluitingscriteria voor opname. Niettegenstaande dit uitsluitingscriterium werden geïnterneerden met comorbide middelhoge en hoge psychopathische trekken aanvaard voor behandeling. In dit hoofdstuk werd de populatie aan de hand van de score op de Psychopathy Checklist-Revised (PCL-R; *N* = 224) verdeeld in drie groepen: laag, gemiddeld en hoge mate van psychopathie. Associaties tussen psychopathie en criminogene risico- en behoeftefactoren werden geanalyseerd. Ook werd de relatie tussen psychopathie en therapiestorend gedrag (gedefinieerd als non-compliance, drop-out en institutioneel wangedrag) onderzocht aan de hand van correlaties en logistische regressie-analyses. De resultaten toonden aan dat psychopathie verband hield met andere risico- en behoeftefactoren, en met therapiestorend gedrag. PCL-R Factor 2 voorspelde institutioneel wangedrag, terwijl PCL-R factor 1 drop-out voorspelde. In zijn geheel bevestigde deze studie ander onderzoek waaruit bleek dat behandeling zich moet richten op criminogene PCL-R Factor 2 kenmerken, maar ook zorgvuldig moet rekening houden met PCL-R Factor 1 kenmerken om ervoor te zorgen dat de behandeling niet afgebroken wordt. Deze studie wees verder op het belang van een responsief behandelingsklimaat om deze moeilijk te behandelen groep in behandeling te houden.

DEEL IV. RISICOTAXATIE TIJDENS EN NA MEDIUM SECURITY BEHANDELING

Deel IV van het proefschrift behandelde een belangrijk onderwerp in de forensische psychiatrie, namelijk de beoordeling van het risico op recidive. Gestructureerde risicotaxatie is een vast onderdeel van de dagelijkse praktijk in forensische instellingen geworden. Tot nog toe werd

echter weinig aandacht besteed aan de klinische toepasbaarheid van de bestaande instrumenten voor risicotaxatie en hun *field validity*.

In **Hoofdstuk 8**, werd de voorspellende waarde van één van de meest gebruikte risicotaxatie-instrumenten, de Historical, Clinical, Risk management-20 (HCR-20) onderzocht. Predictieve validiteit wordt gewoonlijk gemeten aan de hand van de Area Under the Curve (AUC), waarbij de waarschijnlijkheid wordt aangegeven dat een willekeurig gekozen recidivist een hogere score behaalt dan een willekeurig geselecteerde niet-recidivist. Deze retrospectieve maat geeft evenwel geen volledig beeld van de voorspellende waarde van een risicotaxatie-instrument. Het wordt daarom aanbevolen rekening te houden met beide componenten van predictieve validiteit: discriminatie (AUC, sensitiviteit en specificiteit) en calibratie (positief voorspellende waarde, negatief voorspellende waarde, aantal nodig om vast te houden, aantal veilig vrijgesteld). Prospectief-georiënteerde statistieken meten hoe goed de voorspelling van een instrument overeenkomt met de werkelijk waargenomen recidive, en verstrekken bijkomende nuttige informatie omdat zij klinische besluitvorming simuleren. HCR-20 scores gecodeerd binnen één jaar na opname op een medium security unit (*n* = 168) en binnen een jaar voorafgaand aan ontslag uit een medium security unit (*n* = 105) in de periode 2001–2010 werden geanalyseerd in deze studie. Recidive werd op twee momenten geëvalueerd: terwijl de geïnterneerde nog in medium security-behandeling was (verwijzend zowel naar incidenten op de unit als tijdens (on)begeleide vrijheden of ontvluchtingen vanuit de unit), en na ontslag uit de medium security unit. Voorspellingen naar gewelddadige recidive tijdens en na behandeling, gemeten aan de hand van AUC's, bleken niet significant. Uit de analyses bleek dat de HCR-20 vooral nuttig was voor het identificeren van individuen met een laag risico. Mogelijke verklaringen voor de beperkte predictieve validiteit werden in deze studie besproken. Zo kan adequaat risicomanagement bijvoorbeeld de voorspellende nauwkeurigheid beïnvloeden. Uiteindelijk is het belangrijkste doel van risicotaxatie niet het voorspellen van toekomstig geweld, maar de beperking van het risico op geweld.

Hoofdstuk 9 onderzocht de interbeoordelaarsbetrouwbaarheid en predictieve validiteit van de PCL-R wanneer deze afgenomen wordt door beoordelaars in het veld. Internationaal is er een kleine, maar aangroeiende hoeveelheid literatuur die aantoont dat de betrouwbaarheid en predictieve validiteit aanzienlijk verzwakt wanneer de PCL-R in de praktijk wordt toegepast.

Onderzoek gericht op verschillen bij beoordelaars in de praktijk is vooral gericht op twee elkaar niet-uitsluitende verklaringen voor deze verschillen: partijdigheid en individuele verschillen tussen de beoordelaars. De mogelijke schadelijke gevolgen van deze praktijkbevindingen moeten niet licht worden genomen, gezien de bezorgdheid over de stigmatiserende effecten van het label psychopathie op de juridische besluitvorming. Tot nog toe bleef het aantal praktijkstudies (met name met betrekking tot het onderzoek van de PCL-R factorscores) evenwel relatief beperkt, en soms gebaseerd op kleine steekproeven in bijzondere strafrechtsprocedures zoals de *Sexually Violent Predator trials* in de Verenigde Staten. Het is daarom belangrijk te onderzoeken of deze resultaten reproduceerbaar zijn, met name in rechtsgebieden buiten Noord-Amerika en bij andere dan populaties betrokken bij zedendelicten. De huidige studie onderzocht verschillen in scores tussen ziekenhuizen en gevangenissen, maar ook verschillen in predictieve validiteit binnen gelijkaardige instellingen. PCL-R scores werden verzameld, zowel in gevangenis- als in ziekenhuisdossiers, wat resulteerde in 224 PCL-R totaalscores en 74 dubbele scores. Hoewel niet per se partijdigheid in strikte zin werd verwacht, werd toch rekening gehouden met enige contextuele druk over instellingen heen. Bijvoorbeeld, gelet op de druk om geïnterneerden uit de gevangenis te halen, werd verwacht dat beoordelaars in de gevangenis een geïnterneerde lager zouden inschalen op de PCL-R dan een overeenkomstige beoordelaar in een forensische instelling, dit met oog op een mogelijke transfer vanuit de gevangenis. Bij het onderzoek naar de herhaalde metingen werden grote individuele verschillen gevonden samen met een *intraclass* correlatiecoëfficiënt ($ICC_{A,1}$) van .42 voor de PCL-R Totaal score. Afwijkende resultaten werden gevonden voor PCL-R Factor 2, met herhaalde scores binnen dezelfde instelling met een $ICC_{A,1}$ van .28 versus een $ICC_{A,1}$ van 0.57 voor herhaalde scores tussen instellingen. Echter, de AUC voor de PCL-R Totaal- factor- en facetscores verschilde niet significant tussen de instellingen. Voor de volledige onderzoekspopulatie voorspelde PCL-R Factor 2 scores marginaal (gewelddadige) recidive na twee jaar (AUC = .62 en .63), terwijl PCL-R Factor 1 (gewelddadige) recidive niet voorspelde. In overeenstemming met recente praktijkstudies uit andere landen, suggereerden deze resultaten onvoldoende betrouwbaarheid en validiteit in gevangenis- en ziekenhuisinstellingen in Vlaanderen. De klinische implicaties van deze bevindingen werden besproken en omvatten het voorstel om een PCL-R onderzoek enkel te laten uitvoeren door ervaren en getrainde beoordelaars.

DEEL V. RECIDIVE NA MEDIUM SECURITY BEHANDELING

Behandelingsresultaten in de forensische geestelijke gezondheidszorg worden idealiter gemeten over een breed scala van gebieden, met inbegrip van klinische en humanitaire overwegingen. Het voorkomen van toekomstig crimineel gedrag is het belangrijkste doel in de forensische psychiatrische behandeling. Het is daarom verrassend dat er voorafgaand aan het huidige onderzoek geen gegevens over recidive bestonden voor de Vlaamse geïnterneerden. Internationaal varieerden recidivecijfers bij volwassen forensische populaties van 7.1% tot 63.0% voor algemene recidive, en van 1.8% tot 46.0% voor gewelddadige recidive over verschillende follow-up periodes. Intrekkingen van voorwaardelijke vrijstellingen wegens regelovertredingen varieerden van 5% tot 49%.

In **Hoofdstuk 10**, werden recidivecijfers verzameld op basis van het Centraal Strafregister van het ministerie van Justitie. Om het effect van een vrijstelling op proef (VOP) op recidive te onderzoeken, werden strafrechtelijke gegevens van twee subgroepen vergeleken: geïnterneerden onder VOP en geïnterneerden bij wie de internering onvoorwaardelijk werd stopgezet. Ook werd bij de laatste groep een vergelijking gemaakt tussen de VOP-periode en de periode na stopzetting van de internering. De recidivecijfers waren geruststellend laag; 12.5% van de geïnterneerden recidiveerde met een algemeen delict en 7.4% recidiveerde met een geweldsdelict; terwijl na het aflopen van de internering, geen enkele patiënt recidiveerde met een ernstig misdrijf. De studie toonde aan dat er een significante afname was in delictdensiteit voor en na behandeling. Hoewel deze resultaten bemoedigend zijn, blijft het bij gebrek aan een controlegroep onduidelijk of deze effecten kunnen worden toegeschreven aan een behandelingseffect. Bij het vergelijken van de VOP groep met de gedeïnterneerde groep; en bij het vergelijken van VOP-periodes en periodes na stopzetting van de internering bij deze laatste groep, werd duidelijk dat de recidivecijfers het gunstigst waren zolang de interneringsmaatregel van kracht was (dus zolang de patiënt onder toezicht stond). Het verschil was evenwel niet statistisch significant, waarschijnlijk als gevolg van een te laag onderscheidingsvermogen vanwege de kleine aantallen. In dit hoofdstuk werd ook onderzocht hoe vaak VOP's werden ingetrokken. Hieruit bleek dat deze base rates aanzienlijk hoger liggen in vergelijking met recidivecijfers, hetgeen in overeenstemming was met de internationale onderzoeken. Het

intrekken van VOP's, hetgeen plaats vindt in programma's gericht op bescherming van de maatschappij, resulteert in (meer) periodes van opsluiting, en kan worden beschouwd als de keerzijde van de lage recidivecijfers. Patiënten worden terug naar de gevangenis gestuurd, zonder enig formeel proces, hetgeen hun kansen op een nieuwe forensisch psychiatrische opname (of in een later stadium algemeen psychiatrische opname) hypothekeert. Deze heropsluitingen kunnen ook negatieve stereotypen in de reguliere psychiatrie versterken, en het wantrouwen versterken van de forensische patiënten in (forensische) geestelijke gezondheidszorg. Omdat er in de meeste gevallen geen vervolging en veroordeling optreedt, worden de rechten van de potentiële slachtoffers bovendien niet in aanmerking genomen door een strafrechter. Toekomstig onderzoek is nodig om te bepalen of intrekkingen van VOP's kunnen worden aangemerkt als een belangrijke oorzaak van de lage recidivecijfers. Uit internationaal onderzoek is gebleken dat de relatie tussen VOP-intrekkingen en recidives niet lineair was. Over het algemeen suggereerde de huidige studie dat CBM's op een succesvolle manier geïnterneerden (op proef) vrijstelden, hetgeen kan toegewezen worden aan een aantal factoren zoals een gecentraliseerde verantwoordelijkheid, een uniform systeem van toezicht door justitieassistenten, flexibele procedures, en een nauwe samenwerking met de forensische units, wat resulteerde in veilige resocialisatieplannen. Momenteel wordt de nieuwe interneringswet herzien. Verder onderzoek is nodig om de (strafrechtelijke) uitkomsten te evalueren van deze nieuwe wet die meer gebaseerd is op juridische waarborgen met extra bureaucratische procedures en een eerlijk procesverloop.

Hoofdstuk 11 onderzocht recidive in een subpopulatie, namelijk medium security geïnterneerden onder de jurisdictie van CBM Gent (*N* = 192). In het vorige hoofdstuk werden officiële recidivegegevens gekozen als primaire uitkomstmaat voor recidive omdat deze worden beschouwd als de meest betrouwbare bron. Deze uitkomstmaat leidt echter tot onderrapportage, omdat arrestaties en dark numbers niet in rekening worden genomen. Dit kan nog problematischer zijn bij geïnterneerden, omdat werd verwacht dat als misdrijf omschreven incidenten niet automatisch worden vervolgd. Daarom werden in deze studie als misdrijf omschreven incidenten (gedefinieerd als incidenten gecodeerd onder categorieën van het Belgisch wetboek van strafrecht, ongeacht of zij leiden tot verdere vervolging of berechting) die werden gerapporteerd aan de CBM onderzocht met al doel een meer gedetailleerd beeld van

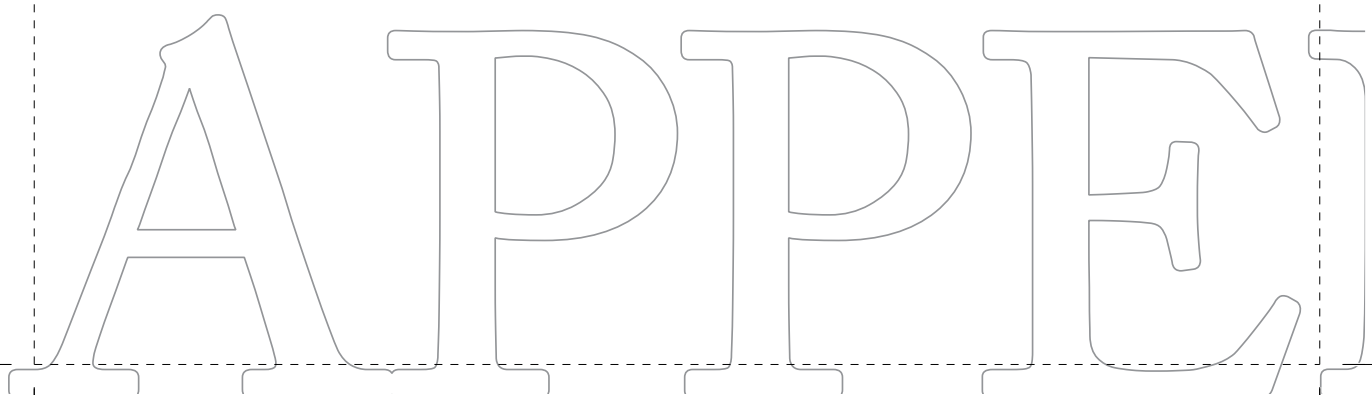
recidive te bieden. Bij significant meer geïnterneerden dan in de officiële recidivecijfers werden incidenten gemeld; de vervijfvoudiging van zowel algemene als gewelddadige recidive was zelfs hoger dan verwacht. Het gebrek aan vervolging kan worden verklaard door het verplichte toezicht van de CBM, dat het de procureur des konings mogelijk maakt een geïnterneerde opnieuw op te sluiten binnen een tamelijk soepele procedure zonder tegensprekelijk debat; bovendien wordt deze procedure onmiddellijk in uitvoering gebracht, waardoor een nieuw langdurig en aanslepend proces vermeden wordt. Zo'n proces zou in veel gevallen toch maar leiden tot een nieuwe interneringsmaatregel. De huidige studie toonde ook aan dat de meeste als misdrijf omschreven feiten minder ernstig van aard waren, wat een andere reden kan zijn dat er niet wordt vervolgd. Naast het onderzoeken van recidive, was het tweede doel van deze studie om risicofactoren voor geweldsrecidive te analyseren. Daartoe werd de groep die hervallen was met een geweldsdelict (n = 43) vergeleken met de groep die niet recidiveerde met een geweldsdelict (n = 102) aan de hand van een aantal risicofactoren voor herval die geïdentificeerd werden in de internationale literatuur. De associatie tussen de geweldsrecidive, en elk van de persoonlijke, delict- en behandelingsgerelateerde kenmerken, werd individueel onderzocht door middel van de Pearson chi-kwadraat en Fisher exact-test voor categorische gegevens en met ongepaarde t-test, of Mann Whitney U-tests voor continue gegevens. Kenmerken die een significante associatie aantoonden met geweldsrecidive werden vervolgens in meerdere logistische regressie-analyses ingevoerd. Verwacht werd dat historische factoren zoals leeftijd bij de eerste veroordeling, of aantal veroordelingen, geweldsrecidive beter zouden voorspellen dan klinische factoren zoals een diagnose van een ernstig psychiatrisch ziektebeeld. Wanneer de kenmerken die geassocieerd waren aan geweldsrecidive werden gecombineerd in één model, was alleen de variabele veelpleger onafhankelijk geassocieerd aan geweldsrecidive. De lage base rate van herval in nieuwe geweldsfeiten was een mogelijke verklaring voor de beperkte risicofactoren die werden teruggevonden. Verder onderzoek met in achtneming van de gehele bevolking werd dan ook aanbevolen.

DEEL VI. SLACHTOFFERS

Dit proefschrift werd afgesloten met een hoofdstuk over de slachtoffers. Uiteindelijk is het doel van forensisch psychiatrische behandeling het voorkomen van nieuwe slachtoffers. Kennis van slachtofferkenmerken (zoals leeftijd, geslacht, en relatie met de dader) maakt een meer effectieve behandeling van geïnterneerden en het voorkomen van toekomstige slachtoffers mogelijk. Verrassend genoeg is onderzoek naar kenmerken van slachtoffers, in tegenstelling tot de wetenschappelijke aandacht voor forensische daders, schaars.

In **Hoofdstuk 12**, werd een overzicht gegeven van de slachtoffers gemaakt bij de index interneringsmaatregel. Informatie over deze delicten werd gehaald uit het Centraal Strafregister van het ministerie van Justitie, en vervolgens vergeleken met informatie beschikbaar bij de diverse parketten. Voor elke geïnterneerde met een geweldsindexdelict werden de volgende slachtofferkenmerken geanalyseerd: aantal slachtoffers, geslacht van het slachtoffer, al dan niet minderjarig zijn van het slachtoffer, en relatie tot het slachtoffer (of het slachtoffer een bekende of een vreemde was voor de dader, dus). Tien deelnemers werden verwijderd uit de analyses omdat alle gegevens over slachtoffers ontbraken, waardoor de analytische studiepopulatie voor deze studie bestond uit 362 geïnterneerden. Herhaalde metingen ANOVA-analyses werden gebruikt voor parametrische gegevens en Wilcoxon signed rank-tests of Friedman-tests werden gebruikt voor niet-parametrische gegevens. De resultaten waren in lijn met eerder onderzoek: volwassenen die bekend waren bij de dader werden het vaakst slachtoffer. Onder de bekende slachtoffers waren familieleden het meest vertegenwoordigd. Vrouwen en mannen hadden een even grote kans om slachtoffer te worden. De studie onderzocht verder of dit slachtofferprofiel (zijnde een volwassene en een kennis) zou verschillen naargelang de psychiatrische diagnose van de dader. De analyses toonden aan dat dit niet het geval was; de slachtofferprofielen waren zeer vergelijkbaar voor daders met verschillende psychiatrische diagnoses. Samengenomen kunnen deze resultaten behulpzaam zijn bij het herkennen van potentieel risicovolle situaties, zowel voor zorgverleners als voor familieleden. Hoewel het geweld gepleegd door patiënten met een ernstig psychiatrisch ziektebeeld in het algemeen een fenomeen is met een lage base rate, kunnen preventieve maatregelen worden genomen op verschillende niveaus, vermits de meest waarschijnlijke slachtoffers bekenden zullen zijn.

De belangrijkste bevindingen, sterktes en zwaktes werden gepresenteerd in een **algemene discussie**. Tot slot werd er een kritische analyse gegeven en aanbevelingen voor toekomstig onderzoek.



Forensische pilootprojecten 'medium security': Incidenten tijdens behandeling van geïnterneerden ressorterend onder CBM Gent

Jeandarme, I., Wittouck, C., Vander Laenen, F., Grouwels, Y., De Varé, J., Oei, T. I., Groenhuijsen, M., & Bogaerts, S.

Panopticon, 36(1), 26-45

ABSTRACT

Research concerning forensic medium security units in Flanders is scarce. This study reports on incidents during treatment. The population consisted of 203 forensic patients conditionally released by the Commission of Social Defense (CPS) of Ghent. During a ten-year follow-up period 236 incident reports were registered, caused by half of the population (53.2%). Less than half of these reports (41.5%) concerned incidents coded as criminal offenses (like drug use, theft or violence). Violent incidents occurred in one out of five reports (20.3%) and were associated with poor treatment compliance and failure. Victims of violence were mainly hospital staff or patients. Incidents were reported to the prosecutor, but rarely led to a new conviction or internment. Instead, one third of the population was (re)incarcerated in prison. Protection, rather than treatment, thus seems to be the primary focus of the CPS.

INLEIDING

Volgens de huidige wet tot Bescherming van de Maatschappij (WBM)[1] kan in België in geval van ontoerekeningsvatbaarheid een interneringsmaatregel van onbepaalde duur worden opgelegd door onderzoeks- en vonnisgerechten. Een interneringsmaatregel is een beveiligingsmaatregel (geen straf) met een tweeledig doel, namelijk de maatschappij beschermen en de geïnterneerde medisch-psychiatrisch behandelen (Cosyns, D'Hont, Janssens, Maes, & Verellen, 2007; Vandevelde et al., 2011). Geïnterneerden vallen onder de bevoegdheid van een Commissie ter Bescherming van de Maatschappij (CBM). In Vlaanderen verblijven geïnterneerden in psychiatrische afdelingen van een gevangenis (*annex*) en in penitentiaire instellingen voor Sociaal Verweer (Merksplas, Turnhout of Brugge), of ze worden behandeld in een forensisch psychiatrische afdeling of in de algemene geestelijke gezondheidszorg (GGZ) (Vander Laenen & De Cauwer, 2011). Bij gebrek aan gepaste doorverwijsmogelijkheden verbleven in 2011 meer dan 600 Vlaamse geïnterneerden in een gevangenis (Moens & Pauwelyn, 2012).

In 2001 werden in drie Vlaamse instellingen forensisch psychiatrische afdelingen opgericht (Openbaar Psychiatrisch Zorgcentrum te Rekem, Sint-Jan-Baptist te Zelzate en Sint-Kamillus te Bierbeek). De afdelingen worden *medium security* afdelingen (MSA) genoemd, refererend naar hun beveiligingsniveau. De focus lag hierbij op geïnterneerden met een gemiddeld veiligheidsrisico voor zichzelf en de samenleving en een gemiddeld risico op herval (Moens & Pauwelyn, 2012). De term *security* verwijst naar een essentieel aspect in de forensische zorg – veiligheid – en bestaat uit drie aspecten: 1) omgevings- of materiele beveiliging, 2) procedurele beveiliging en 3) relationele beveiliging[2] (Kennedy, 2002). Waar de twee eerste componenten uitgaan van het aanbodperspectief, is de derde component nauw verbonden met het risicoprofiel en de veiligheidsnoden van de patiënt (H. Vertommen, persoonlijke communicatie, 13 augustus, 2013). Het MSA zorgaanbod is bedoeld voor Nederlandstalige geïnterneerden met enige mate van motivatie en leerbaarheid en, op psychiatrisch vlak, bij voorkeur met een psychotische en/of een persoonlijkheidsstoornis. Exclusiecriteria voor opname zijn een op de voorgrond staande seksuele[3] of verslavingsproblematiek en/of een hoge mate van psychopathie (De Smedt, Mariën, & Vermeiren, 2008). Alleen de MSA van Zelzate heeft sinds 2006 een aanbod voor vrouwelijke geïnterneerden.

De algemene GGZ is vaak terughoudend tegenover de opvang en behandeling van daders met een psychiatrische problematiek, niet in het minst omdat deze patiëntengroep *de reputatie heeft* therapieresistent, moeilijk behandelbaar, weinig therapietrouw en gewelddadig te zijn (Lamb, Weinberger, & Gross, 1999; Schanda, Stompe, & Ortwein-Swoboda, 2009). In een onderzoek van 1994 naar behandelmogelijkheden van geïnterneerden in Vlaamse psychiatrische ziekenhuizen, bleek er terughoudendheid om teveel geïnterneerden op te nemen onder meer vanuit een bezorgdheid om de veiligheid voor het personeel en het imago van de instellingen. Ook de behandelmotivatie van geïnterneerden werd in vraag gesteld. Tenslotte bleek er behoefte aan sterker beveiligde afdelingen met betere omkadering dan op dat moment beschikbaar was in de GGZ (Cosyns, Dillen, De Ruyter, & De Doncker, 1994). Vooral de idee dat regelmatig geweldsincidenten voorkomen binnen deze populatie wordt als problematisch beschouwd. Toch moet worden vastgesteld dat prevalentiecijfers van geweldsincidenten bij deze populatie in Vlaanderen tot nog toe ontbreken. Dit alles neemt niet weg dat, wanneer zich geweldsincidenten voordoen in een voorziening, deze een invloed hebben op de stabiliteit van een voorziening, zorgen voor een groot personeelsverloop en in negatieve zin het behandelproces beïnvloeden (Gow, Choo, Darjee, Gould, & Steele, 2010). In dit artikel worden de prevalentie en kenmerken van incidenten tijdens MSA opname bestudeerd, naast kenmerken van de populatie die incidenten veroorzaakt. Niet enkel kenmerken van de populatie, maar ook situationele factoren zoals overbevolking, de mate en kwaliteit van interacties tussen patiënten en personeel en het agressiebeleid binnen een institutie spelen een rol bij het tot stand komen van een geweldsincident (Gadon, Johnstone, & Cooke, 2006). Het huidig onderzoeksopzet liet echter niet toe deze andere factoren te bestuderen.

Uit Nederlands onderzoek in de reguliere GGZ blijkt dat geweld op de werkvloer zelden leidt tot een strafrechtelijke vervolging. Slechts een vierde van de (pogingen tot) fysieke geweldsincidenten werden gerapporteerd bij de politie, en ongeveer 10% van de gerapporteerde incidenten werd voor de rechter gebracht (Harte, Van Leeuwen, & Theuws, 2013). De huidige studie is voor zover ons bekend de eerste die de strafrechtelijke reactie op incidentmeldingen bij een forensische populatie in kaart brengt.

Internationaal onderzoek naar incidenten in MSA

De meest voorkomende incidenten die in de internationale literatuur worden beschreven en die toepasbaar zijn op *incidenten* in MSA zijn op de eerste plaats zelfbeschadigend gedrag, verbaal geweld en fysiek geweld. Ook het gebruik van alcohol of illegale drugs en ontvluchting worden gerapporteerd (Abidin et al., 2013; Blattner & Dolan, 2009; Gow et al., 2010; Hillbrand, 1995). Een ontvluchting blijkt zelden gepaard te gaan met het plegen van een misdrijf (Gradillas, Williams, Walsch, & Fahy, 2007). Seksueel geweld en brandstichting komen minder vaak voor (Blattner & Dolan, 2009; Gow et al., 2010). Meestal is het personeel het slachtoffer van fysiek geweld, pas in tweede instantie onbekenden of medepatiënten (Gow et al., 2010; Gradillas et al., 2007). Wat *persoonskenmerken* betreft van patiënten die betrokken zijn bij geweldsincidenten, wordt een persoonlijkheidsstoornis geassocieerd met fysiek geweld en in veel mindere mate met een majeur psychiatrisch ziektebeeld, zoals schizofrenie (Gow et al., 2010). Een hoge score op de Psychopathy Checklist-Revised (PCL-R; Hare, 2003) wordt intramuraal meer geassocieerd met algemene incidenten dan met gewelddadige incidenten (Guy, Edens, Anthony, & Douglas, 2005). Eerder onderzoek toonde aan dat een verhoogde score op het risicotaxatie-instrument Historical, Clinical, Risk Management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997)) gerelateerd is aan intramurale geweldsincidenten (McDermott, Edens, Quanbeck, Busse, & Scott, 2008; Mudde, Nijman, Van der Hulst, & Van den Bout, 2011).

Internationale vergelijkingen van wetenschappelijke onderzoeksbevindingen dienen steeds met de nodige voorzichtigheid te gebeuren. Verschillen in onder andere therapeutische omkadering, wetgeving en beleid rond internering kunnen immers een invloed hebben op deze bevindingen. De vergelijking van de kenmerken van de MSA patiëntenpopulatie in bovenstaande studies uit Ierland (Abidin et al., 2013), het Verenigd Koninkrijk (Blattner & Dolan, 2009; Gow et al., 2010; Gradillas et al., 2007), de staat Californië (McDermott et al., 2008) en Nederland (Mudde et al., 2011), met de kenmerken van de patiëntenpopulatie in onderhavige MSA studie geeft aan dat deze gelijkaardig zijn voor demografische en juridische kenmerken. Ze zijn minder vergelijkbaar voor klinische kenmerken aangezien MSA patiënten in internationale studies vaker gediagnosticeerd worden met een psychotische problematiek.

De huidige studie

Dit onderzoek maakt onderdeel uit van een grotere studie[4] die werd opgestart bij de oprichting van de MSA-afdelingen en wordt gefinancierd door het OPZC Rekem, *Limburg Sterk Merk* en *Samenwerking Psychiatrische Initiatieven Limburg*. De studie werd goedgekeurd door het ethisch comité van het Universitair Ziekenhuis Antwerpen op 24 januari 2011. In deze studie werden incidentmeldingen tijdens MSA behandeling van een subgroep geïnterneerden onderzocht. Hierbij werden zowel gebeurtenisgerelateerde kenmerken – zoals het profiel van slachtoffers en de strafrechtelijke reactie – als persoonsgerelateerde kenmerken – zoals het profiel van de geïnterneerde – van de incidenten bestudeerd. De onderzoeksdoelstelling werd vertaald in de volgende onderzoeksvragen:

- Wat was de frequentie, aard en ernst van de geregistreerde incidenten?
- Hoeveel slachtoffers waren betrokken bij de incidenten en wat was hun profiel?
- Hoe werd strafrechtelijk gereageerd op de incidentmeldingen?
- Verschilde de populatie met geweldsincidenten van de populatie zonder geweldsincidenten?

METHODE

Dataverzameling

Tussen 2001 en 2010 werden de dossiers van 203 geïnterneerden die behandeld werden in een van de drie MSA's (Zelzate: 63.1%, *n* = 128; Bierbeek: 21.2%, *n* = 43 en Rekem: 15.8%, *n* = 32) retrospectief geanalyseerd.[5] In het huidige onderzoek werden enkel geïnterneerden, ressorterend onder CBM Gent geïncludeerd. De steekproef bedraagt 37.5% van het totaal aantal geïnterneerden in deze periode behandeld in Vlaamse MSA's en 98.5% van de populatie geïnterneerden door CBM Gent verwezen naar medium security behandeling.

Voor dit onderzoek werd gebruik gemaakt van verschillende informatiebronnen. Sinds de opstart van de MSA's verzamelen de behandelaars op systematische wijze demografische, klinische en juridische gegevens betreffende antecedenten en indexdelict[6] van de opgenomen geïnterneerden. Deze gegevens werden door de onderzoekers retrospectief (periode 2011–2013)

gecontroleerd. Teneinde een zo valide mogelijke scoring van de variabelen te bekomen, werd een checklist en bijhorende handleiding opgesteld. Indien informatie ontbrak, werd het behandeldossier van de geïnterneerde ingezien. Daarnaast werden alle CBM dossiers van de geïnterneerden doorgenomen teneinde justitiële antecedenten, incidentmeldingen en CBM beslissingen te verzamelen. De antecedenten en indexdelicten werden geclusterd in vijf categorieën van delicten: gewelds delicten, eigendomsdelicten, drugsdelicten, seksuele delicten en overige delicten. Om deze clustering mogelijk te maken, werd gebruik gemaakt van een codeboek, waarbij aan elke code uit het strafwetboek een categorie werd toegekend.

Informatie betreffende de kenmerken van de geïnterneerden (persoonsgerelateerd) werd verzameld via CBM dossiers en MSA behandeldossiers.

Informatie over incidenten tijdens opname (gebeurtenisgerelateerd) werd verzameld via CBM dossiers van geïnterneerden die 1) onder de bevoegdheid vielen van de Gentse CBM en 2) die tussen 2001 en 2010 werden opgenomen. Een incident werd gedefinieerd als een (negatieve) gebeurtenis of een cluster van (negatieve) gebeurtenissen (incidentmelding) die door justitieassistenten of behandelaren gerapporteerd werden aan de CBM. Het gaat hierbij om incidenten waarvan de kwalificatie in het strafwetboek is terug te vinden (als misdrijf omschreven incident) en om incidenten die te maken hebben met het overtreden van de interneringsvoorwaarden (niet als misdrijf omschreven incident). Voor elk incident werd de aard en ernst onderzocht (Tabel 1), werd nagegaan of er slachtoffers betrokken waren en werd de aard van de relatie tussen de geïnterneerde en het slachtoffer onderzocht. Alle slachtoffers van geweldsincidenten waarbij de lichamelijke integriteit werd aangetast of ernstig bedreigd, werden geregistreerd.

De strafrechtelijke reactie op de incidenten werd nagegaan via 1) de beslissingen van de CBM, 2) de detentiefiches opgevraagd via het administratief registratiesysteem *Detentie Informatie Systeem* (SIDIS) van het Directoraat-generaal Penitentiaire Inrichtingen van het gevangeniswezen[7] en 3) de uittreksels van de strafregisters opgevraagd via het Centraal Strafregister van de Federale Overheidsdienst Justitie.[8] In de CBM dossiers werd nagegaan of naar aanleiding van een incidentmelding naast een eventuele oproeping ter vermaning ook strafrechtelijke actie werd ondernomen, dan wel of de interneringsvoorwaarden werden

Tabel 1. Codering incidenten[9]

Als misdrijf omschreven incident	
(Seksueel) gewelddadig incident	Het toebrengen van letsel aan één of meerdere personen, dan wel een poging daartoe of dreigen ermee (Webster et al., 1997)
Levensdelict	
Poging levensdelict	
Brandstichting met gevaar voor personen	
Overig geweldsdelict	
Eigendomsdelict met geweld	
Verbaal geweldsdelict	
Seksueel geweldsdelict tegen minderjarigen	
Seksueel geweldsdelict tegen meerderjarigen	
Niet-gewelddadig seksueel delict	Seksueel hands-off delict
Eigendomsdelict	
Eigendomsdelict zonder geweld (bijvoorbeeld, diefstal, oplichting)	
Brandstichting zonder gevaar voor personen	
Drugsdelict	Bezit en verkoop van illegale drugs en niet-voorgeschreven medicatie
Overig delict	
Delict in de familiale sfeer bijvoorbeeld, familieverlating, illegale abortus	
Verkeersinbreuk	
Ander delict bijvoorbeeld vernieling, verstoring openbare orde	
Niet als misdrijf omschreven incident	
Compliance	Het niet opvolgen van behandelafspraken (bijvoorbeeld zich onthouden van alcoholgebruik, geen relatie aangaan met een medepatiënt, regelmatige therapiedeelname, inname van voorgeschreven medicatie)
Schending overige voorwaarden	Het niet opvolgen van bijkomende juridische voorwaarden opgelegd door de CBM (bijvoorbeeld niet nakomen afspraken justitie assistent of verbreken contactverbod)
Ongeoorloofde afwezigheid	
Onttrekking	Ongeoorloofde afwezigheid tijdens begeleid of onbegeleid verlof
Ontvluchting	Ongeoorloofde afwezigheid zonder voorafgaandelijk toegestaan verlof

aangepast door de CBM of werd overgegaan tot een (her)opsluiting in de gevangenis. Bij een (her)opsluiting werd de duur van de detentie berekend. Indien een incident door het openbaar ministerie vervolgd werd en uiteindelijk gevonnist, werd de aard van het vonnis bepaald.

Beschrijving van de steekproef

De gemiddelde leeftijd op het moment van opname was 35.7 jaar (*SD* = 10.23, range = 18.8–73.4). De respondenten waren overwegend mannen[10] (92.6%, *n* = 188). De meeste geïnterneerden waren in België geboren[11] (92%, *n* = 183). Ongeveer vier op tien geïnterneerden (42.4%, *n* = 86) werd meer dan een keer opgenomen op een MSA. De gemiddelde opnameduur bedroeg gecumuleerd[12] 627.3 dagen (*SD* = 466.2 dagen; range = 8–2729 dagen). Twee geïnterneerden overleden in de loop van de behandeling en voor 17 geïnterneerden was bij het afsluiten van de studie geen ontslagdatum bekend. In de groep van de beëindigde behandelingen (*n* = 184) werd ongeveer een kwart (26.6%, *n* = 49) vroegtijdig afgesloten (drop-out behandelingen). Deze drop-out behandelingen hebben zowel betrekking op behandelingen door de geïnterneerde zelf afgebroken, tegen advies van de MSA, als op behandelingen die vroegtijdig stopgezet werden door het ziekenhuis wegens conflicten in de samenwerking of wegens incidenten.

De indexinternering werd vooral opgelegd naar aanleiding van gewelds delicten (70.4%, *n* = 143), gevolgd door eigendomsdelicten (22.7%, *n* = 46), en in beperkte mate voor drugsdelicten (2%, *n* = 4), seksuele delicten inclusief seksueel gewelddadige delicten (4.4%, *n* = 9) en overige delicten (0.5%, *n* = 1). In geval van meerdere misdrijven werd het indexdelict gekwalificeerd aan de hand van het meest ernstige delict per indexinternering. De ernst van het delict werd gecodeerd naargelang morele ernst, grotendeels gebaseerd op de Delict Ernst score (DE-12; Brand, 2005). Een minderheid (14.3%, *n* = 29) had een blanco strafregister voorafgaand aan de indexinternering. De gemiddelde leeftijd bij de eerste veroordeling of interneringsmaatregel was 25.2 jaar (*SD* = 9.5; range = 9.8–66.5). Het merendeel van de steekproef (94.1%, *n* = 191) was eerder veroordeeld of geïnterneerd wegens (seksuele) gewelds delicten. Meer dan een derde kon als veelpleger beschouwd worden (33.9%, *n* = 57).[13] Bij 16.3% van de geïnterneerden (*n* = 33) was er voorafgaand aan het indexdelict een zeer ernstig persoonsdelict gedefinieerd, met name (poging tot) moord/doodslag, brandstichting met gevaar voor personen, (poging tot) verkrachting en (poging tot) aanranding van de eerbaarheid van minderjarigen.[14]

Diagnoses op basis van de DSM-IV(-TR) (American Psychiatric Association, 2000) werden door de eerste auteur in overleg met de behandelende psychiaters gesteld op basis van alle beschikbare dossierinformatie. Voor deze studie werden de diagnoses geclusterd in drie grote categorieën: a) majeure psychiatrische stoornissen (MPS) (37.4%, *n* = 76); b) majeure psychiatrische stoornissen in combinatie met aan middelen gebonden stoornissen[15] (MPS-MGS) (19.2%, *n* = 39) inclusief direct hieraan gerelateerde complicaties; en c) overige stoornissen (AN) (43.4%, *n* = 88) waaronder – al dan niet gecombineerd – persoonlijkheidsstoornissen (*n* = 74), middelen gerelateerde stoornissen (*n* = 55), zwakbegaafdheid of verstandelijke beperking (*n* = 29) en alle andere diagnoses (*n* = 14) (Monahan et al., 2001). Met majeure psychiatrische stoornissen worden ernstige psychiatrische ziektes zoals schizofrenie en bipolaire stoornis bedoeld. Onder middelen gerelateerde stoornissen worden zowel problematisch alcohol- en/of illegaal druggebruik als afhankelijkheid aan deze middelen verstaan. Zwakbegaafdheid refereert naar een verlaagd intelligentiequotiënt tussen 70 en 84; verstandelijke beperking refereert naar personen met een IQ beneden de 70 die al voor hun 18de levensjaar op verschillende levensdomeinen problemen ondervonden (American Psychiatric Association, 2000; Schalock et al., 2010). Andere diagnoses verwijzen naar minder ernstige psychiatrische ziektebeelden zoals aanpassingsstoornissen.

Van een aanzienlijk deel van de populatie (42.4%, *n* = 86) was geen IQ-score beschikbaar op basis van de Wechsler Adult Intelligence Scale-III (WAIS-III; Wechsler, 2000). Bij de personen waarvan wel een score bekend was, bedroeg deze gemiddeld 78.8 (*SD* = 18.11; range = 48–135). Ook risicotaxatie werd niet systematisch uitgevoerd: 129 scores op de PCL-R (63.6%) en 78 scores op de HCR-20 (38.4%) ontbraken. Meer dan een vierde (28.8%, *n* = 36) werd geclassificeerd als hoog risico op basis van de HCR-20.[16] De gemiddelde PCL-R score bedroeg 20.6 (*SD* = 6.5; range = 8–34). Meer dan een kwart scoorde boven de cut-off van 25 (29.2%, *n* = 21).[17]

Samenvattend kan gesteld worden dat de MSA populatie wordt gekenmerkt door een complexe problematiek zowel op sociaal, juridisch als psychiatrisch vlak. Wat betreft demografische en juridische gegevens is de steekproef vergelijkbaar met forensisch psychiatrische populaties in het buitenland. Wat betreft klinische kenmerken valt de hoge prevalentie van psychotische stoornissen in buitenlandse studies op, in vergelijking met onze MSA-studie. De drop-out subgroep verschilde op een aantal kenmerken van de groep

geïnterneerden die de behandeling wel afmaakte. Gemiddeld was deze groep jonger bij eerste opname en waren er meer personen met een zeer ernstig persoonsdelict in de voorgeschiedenis ($\chi^2(1) = 7.37$, $p = .007$). Ook was de gecumuleerde opnameduur gemiddeld korter ($U = 1838.0$, $z = -4.60$, $p = .00$) en verschilde de psychiatrische problematiek ($\chi^2(2) = 6.27$, $p = .04$).

ONDERZOEKSRESULTATEN[18]

Gebeurtenisgerelateerde kenmerken van de incidenten

In totaal werden tussen 2001 en 2010 236 incidentmeldingen – bestaande uit een of meerdere incidenten die in dezelfde tijdspanne plaatsvonden – aan de CBM gerapporteerd (Tabel 2). Het totaal aantal incidenten bedroeg 331.

Aard en ernst van de incidenten

Bijna twee derde (62.6%, *n* = 206) van de incidenten betrof niet als misdrijf omschreven incidenten.[19] Meer specifiek ging het om ongeoorloofde afwezigheden (*n* = 122), problemen met het opvolgen van behandelafspraken (compliance; *n* = 71) en overige schendingen van voorwaarden (*n* = 13). Ongeoorloofde afwezigheden bestonden uit ontvluchtingen uit het ziekenhuis(terrein) (70.5%, *n* = 86), of onttrekkingen aan een toegekend begeleid verlof (27%, *n* = 33), of onbegeleid verlof (2.5%, *n* = 3). De gemiddelde duur van een ongeoorloofde afwezigheid bedroeg 6.38 dagen (*SD* = 14.7, range = 1–87).[20] In 71.2% van de gevallen werden de geïnterneerden binnen twee dagen terug gevat. Een vijfde van de ontvluchtingen (20.5%, *n* = 25) ging gepaard met een of meer potentieel strafbare incidenten (cf. infra).

Iets meer dan een derde van de incidenten (37.4%, *n* = 123) konden als misdrijf omschreven worden. In ongeveer de helft van de gevallen (48.8%, *n* = 60) betrof het niet-gewelddadige feiten: drugbezit en/of -verkoop (*n* = 34), eigendomsdelicten zonder geweld (*n* = 17), overige delicten zoals vernielingen (*n* = 8) en hands-off seksuele delicten (*n* = 1). In de andere helft (51.2%, *n* = 63) ging het om gewelddadige feiten: vooral verbaal geweld (zoals bedreigingen en stalking) (*n* = 29) en slagen en verwondingen (*n* = 28). Eigendomsdelicten met geweld (zoals handtassendiefstal) (*n* = 3), pogingen tot een levensdelict (*n* = 2) en brandstichting met gevaar voor personen (*n* = 1) kwamen uitzonderlijk voor.

Tabel 2. Aard van incidenten (gebeurtenis-gerelateerd)

	Incidenten event based (*n* = 329)
Niet als misdrijf omschreven incidenten	206 (62.6%)
Ongeoorloofde afwezigheid	122 (37.1%)
Compliance	71 (21.6%)
Schending strafrechtelijke voorwaarden	13 (4.0%)
Als misdrijf omschreven incidenten	123 (37.4%)
Geweldsdelicten	63 (19.1%)
Verbaal geweldsdelict	29
Overig geweldsdelict	28
Eigendomsdelict met geweld	3
Poging levensdelict	2
Brandstichting met gevaar voor personen	1
Levensdelict	0
Seksueel hands-on delict	0
Drugsdelicten	34 (10.3%)
Cannabis	16
Niet anders omschreven	10
Heroïne	3
Illegale medicatie	3
Speed, cocaïne	2
Hallucinogene paddenstoelen	1
Eigendomsdelicten	17 (5.2%)
Eigendomsdelict zonder geweld	15
Brandstichting zonder gevaar voor personen	2
Overige delicten	8 (2.4%)
Ander delict	8
Delict in de familiale sfeer	0
Verkeersinbreuk	0
Seksuele delicten	1 (0.3%)
Seksueel hands-off delict	1
Totaal aantal incidenten	329 (100%)

Geweldsincidenten vonden zelden plaats in samenhang met drugbezit (4.8%, *n* = 3, *p* = .44) en in 14.3% van de gevallen (*n* = 9) met alcoholgebruik (RR = 2.9; 95% BI = 1.6–5.2; *p* = .08). Een kwart van de als misdrijf omschreven incidenten vond plaats tijdens een ongeoorloofde afwezigheid

(24.4%, *n* = 30). Meer specifiek betrof het hier drugbezit (*n* = 12), diefstal (*n* = 8), bedreiging (*n* = 6), fysieke agressie (*n* = 3) en vernieling (*n* = 1).

Slachtoffergegevens

Bij een vijfde van de incidentmeldingen (20.3%, *n* = 48) waren een of meerdere directe slachtoffers betrokken. In totaal ging het om 64 slachtoffers van geweldsincidenten; geen enkel slachtoffer was minderjarig. Overwegend mannen waren slachtoffer (80.0%, *n* = 20).[21] In de meeste gevallen waren dader en slachtoffer bekenden van elkaar (93.3%, *n* = 56); hulpverleners[22] (45.0%, *n* = 27), medepatiënten (28.3%, *n* = 17) en familieleden (11.7%, *n* = 7) waren het vaakst slachtoffer.

Strafrechtelijke afhandeling

De incidentmeldingen – 236 in totaal – bestonden uit niet als strafbaar omschreven (*n* = 138) en als strafbaar omschreven incidenten (*n* = 98). In Tabel 3 wordt een overzicht gegeven van de beslissing van de CBM per type incidentmelding, gekwalificeerd aan de hand van het meest ernstige incident per melding.[23] In 54.3% (*n* = 127) van de meldingen besliste de CBM tot verdere opname in de MSA zonder (strafrechtelijk) gevolg[24], in 44.4% (*n* = 104) van de meldingen tot een (her)opsluiting in de gevangenis en in 1.3% (*n* = 3) van de meldingen tot een aanpassing van de voorwaarden (zoals het bijkomend opleggen van een contactverbod).

Wanneer werd beslist tot een (her)opsluiting in de gevangenis was de aanleiding in 37 gevallen (35.6%) een ongeoorloofde afwezigheid, in 28 gevallen (26.9%) een melding van geweld, in 16 gevallen (15.4%) compliance problemen, in 11 gevallen (10.6%) drugsbezit, in 9 gevallen (8.7%) eigendomsdelicten en in 3 gevallen (3.9%) overige incidenten. Opvallend is dat meldingen van als strafbaar omschreven incidenten niet vaker (52.0%, *n* = 51) dan meldingen van niet als strafbaar omschreven incidenten leidden tot een aanhouding (39.0%, *n* = 53), hoewel een mogelijk verband gesuggereerd werd (*p* = .06). Bovendien bleek geen significant verband tussen het risico op een (her)opsluiting in de gevangenis en de aard van de incidentmelding (*p* = .38). De detentietijd volgend op een (her)opsluiting duurde gemiddeld 443.2 dagen (*SD* = 461.6, range = 3–1912).[25] Slechts bij twee meldingen – in beide gevallen een ontvluchting met ernstige

geweldsdelicten – volgde parallel aan de indexinternering een nieuw strafrechtelijke vonnis, waarvan een interneringsmaatregel en een veroordeling in Nederland tot een TBS-maatregel.

Tabel 3. Beslissing van de Commissies ter Bescherming van de Maatschappij in functie van het type incidentmelding

	Meldingen totaal (*n* = 234)	Onttrekking (*n* = 108)	Geweld en seksueel (*n* = 48)		Drugs (*n* = 31)	Compliance (*n* = 28)	Overig delict (*n* = 20)	
			Geweld (*n* = 47)	Seksueel hands-off (*n* = 1)			Eigendom (*n* = 13)	Ander (*n* = 6)
Heraanhouding	104 (44.4%)	37 (35.6 %)	28 (26.9 %)	0	11 (10.6 %)	16 (15.4 %)	9 (8.7 %)	3 (3.9 %)
Aanpassing voorwaarden	3 (1.3%)	0	1	0	0	1	0	1
Melding dossier	127 (54.3%)	71	18	1	20	11	4	2

Persoonsgerelateerde kenmerken van de incidenten

In totaal waren iets meer dan de helft van de geïnterneerden (53.2%, *n* = 108) betrokken bij de incidentenmeldingen. Gemiddeld genomen was een incidentpleger bij 2.2 meldingen betrokken (*SD* = 1.7; range = 1–9).

Bijna de helft (42.9%, *n* = 87) van de geïnterneerden was betrokken bij een incidentmelding zonder strafbare feiten. Ongeveer een derde van de geïnterneerden (32.0%, *n* = 65) was ongeoorloofd afwezig en een kwart (23.2%, *n* = 47) hield zich niet aan de behandelafspraken. Slechts een minderheid (2.5%, *n* = 5) schond overige juridische voorwaarden.

Een derde van de geïnterneerden (31.0%, *n* = 63) was betrokken bij een incidentmelding met strafbare feiten. Minder dan een op vijf geïnterneerden (17.7%, *n* = 36) was betrokken bij een geweldsdelict. Daarnaast pleegde 10.3% van de populatie (*n* = 21) een drugsdelict, 7.4% (*n* = 15) een eigendomsdelict en 3.9% (*n* = 8) een overig delict, veelal vernielingen. Een geïnterneerde (0.5%) pleegde een seksueel hands-off delict; geen enkele een seksueel hands-on delict.

Een derde van de populatie (34.5%, *n* =70) werd in de loop van de MSA behandeling (opnieuw) opgesloten in de gevangenis naar aanleiding van een incident.

Kenmerken van geïnterneerden met geweldsincident en geïnterneerden zonder geweldsincident Tabel 4)

Demografische kenmerken

Zoals kon verwacht worden, bestond de groep geïnterneerden die een geweldsincident veroorzaakte overwegend uit mannen (94.6%) met de Belgische nationaliteit (94.6%). Hun gemiddelde leeftijd bij opname was 35.4 jaar (*SD* = 9.8, range = 23.2–73.4). Er werd geen verband geobserveerd tussen de demografische variabelen en het veroorzaken van een geweldsincident.

Justitiële antecedenten

De groep geïnterneerden die een geweldsincident veroorzaakte bestond voor 34.6% uit veelplegers, had in 97.3% van de gevallen een (seksueel) geweldsincident in de voorgeschiedenis en in 24.3% een zeer ernstig antecedent in de voorgeschiedenis. De gemiddelde leeftijd op moment van het eerste vonnis was 25.4 jaar (*SD* = 9.8, range = 15.4-66.5). Het indexdelict bestond uit geweldsdelicten (78.4%), eigendomsdelicten (18.9%) en seksuele delicten (2.7%). Er werd geen verband geobserveerd tussen de justitiële antecedenten en het veroorzaken van een geweldsincident.

Klinische kenmerken

In de groep geïnterneerden die een geweldsincident veroorzaakte, had 29.7% een diagnose van een majeure psychiatrische stoornis, 10.8% van een majeure psychiatrische stoornis in combinatie met een aan een middelen gebonden stoornis en 59.5% een andere stoornis. Drie kwart van de groep had een diagnose van een persoonlijkheidsstoornis (75.7%) en was normaal begaafd (78.4%). Een vierde van de onderzoeksubjecten had een hoge PCL-R score (25.0%) en bijna de helft scoorde hoog op de HCR-20 (45.8%). Het veroorzaken van een geweldsincident hing positief samen met de score hoog op de HCR-20 (RR = 1.9; 95% BI = 1.1–3.2*; p* = .05).

Tabel 4. Persoonsgerelateerde kenmerken van de incidenten

	Totale populatie (*n* = 203)		Groep met geweldsincidenten (*n* = 37)		Groep zonder geweldsincidenten (*n* = 166)	
	n	%	*n*	%	*n*	%
Categoriale variabelen						
Demografische variabelen						
Mannelijk geslacht	188	92.61	35	94.59	153	92.17
In België geboren[a]	183	91.96	35	94.59	148	91.36
Justitiële antecedenten						
Veelpleger	57	33.93	13	3.8.24	44	32.84
Zeer ernstig antecedent	33	16.26	9	24.32	24	14.46
Voorgeschiedenis (seksueel) geweld	191	94.09	36	97.30	155	93.37
Indexdelict						
Geweld	143	70.44	29	78.38	114	68.67
Eigendom	46	22.66	7	18.92	39	23.49
Drugs	4	1.97	0	0	4	2.41
Seksueel	9	4.43	1	2.70	8	4.82
Overig	1	0.49	0	0	1	0.6
Behandelingskenmerken						
Ongeoorloofde afwezigheid	72	35.47	17	45.95	55	33.13
Compliance	81	39.90	26**	70.27	55**	33.13
Drop out	49	26.63	17**	50	32**	21.33
Drugbezit/-handel	21	10.35	6	16.22	15	9.04
Klinische kenmerken						
PCL-R ≥ 25[b]	21	29.17	4	25	17	30.36
HCR-20 hoog[c]	36	28.80	11*	45.83	25*	24.75
Psychiatrische diagnose						
MPS	76	37.44	11	29.73	65	39.16
MPS-MGS	39	19.21	4	10.81	35	21.08
AN	88	43.35	22	59.46	66	39.76
Persoonlijkheidsstoornis	127	62.56	28	75.68	99	59.64
Middelenstoornis	94	46.31	17	45.95	77	46.39
Intellectuele beperkingen						
Normaal begaafd	157	77.34	29	78.38	128	77.11
Zwakbegaafd	13	6.40	0	0	13	7.83
Verstandelijk beperkt	33	16.26	8	21.62	25	15.06
Continue variabelen	*M*	*SD*	*M*	*SD*	*M*	*SD*
Demografische variabelen						
Leeftijd bij eerste opname	35.66	10.23	35.38	9.81	35.72	10.35
Justitiële antecedenten						
Leeftijd eerste vonnis	25.17	9.54	25.41	9.78	24.82	9.71

Noot. MPS = majeure psychiatrische stoornis; MPS-MGS = majeure psychiatrische stoornis in combinatie met een aan middelen gebonden stoornis; AN = overige stoornis.
[a] 2.0% missing; [b] 55.0% missing; [c] 38.4% missing.
* $p < .05$. ** $p < .01$.

Behandelkenmerken

Ongeveer de helft van de groep geïnterneerden die een geweldsincident veroorzaakte, was minstens eenmaal ongeoorloofd afwezig tijdens MSA opname (46.0%) en/of maakte de behandeling niet af (50%). Daarnaast vertoonde 70.3% een gebrekkige compliance. Het veroorzaken van een geweldsincident was geassocieerd met drop-out behandelingen (RR = 2.8; 95% BI = 1.5–5; p = .00), en een gebrekkige compliance (RR = 3.6; 95% BI = 1.9–6.8; p = .00). Alcoholgebruik als onderdeel van gebrekkige compliance hing sterk positief samen met het plegen van geweldsincidenten (p = .00).

Tot slot werden de individuele significante associaties met geweldsincidenten opgenomen in een meervoudige logistische regressie. Hier bleek de associatie met de HCR-20 score niet langer significant. Een gebrekkige compliance en het niet afmaken van de behandeling hing samen met een verhoogd risico op geweldsincidenten (respectievelijk, p = .01 en p = .02). Corrigerend voor het al dan niet afmaken van de behandeling, was er bij gebrekkige compliance meer kans op geweldsincidenten dan bij goede compliance (RR = 2.2; 95% BI = 1.1–3.7). Corrigerend voor compliance was er meer kans op geweldsincidenten bij het niet afmaken van de behandeling dan wanneer de behandeling voltooid werd (RR = 2.5; 95% BI = 1.3–4.5).

DISCUSSIE

In dit artikel werden de kenmerken van incidenten die voorkwamen tijdens opname in drie forensisch psychiatrische afdelingen voor medium security (MSA) en de kenmerken van de geïnterneerden die bij deze incidenten betrokken waren bestudeerd. De meerderheid van de incidenten zijn niet strafbaar en niet gewelddadig. De helft van de geïnterneerden die werden opgenomen op een MSA tijdens de observatieperiode van tien jaar was betrokken bij een incident. Belangrijk evenwel is dat er in twee derde van de incidenten geen sprake was van een als misdrijf omschreven feit, maar dat het ging om regelovertredingen zoals problemen in het behandelingsverloop, met name ontvluchtingen en het niet nakomen van behandelafspraken.

In vergelijking met internationale cijfergegevens (5.5%; Blattner & Dolan, 2009; 8.3%; Gow et al., 2010), ligt het aantal ontvluchtingen in ons onderzoek hoog (37.1%). Een mogelijke verklaring hiervoor is dat medium security geïnterneerden relatief snel na hun opname

(on)begeleide vrijheden kregen, bijvoorbeeld met het oog op therapiedeelname op het ziekenhuisterrein. Vrijheden worden progressief aan de hand van risicotaxaties toegekend, maar houden een berekend risico in. Architecturaal is de veiligheidsgraad van een MSA, die is ingebed in de setting van een regulier psychiatrisch ziekenhuis, niet te vergelijken met een gevangenis of het nieuw gebouwde forensisch psychiatrisch centrum in Gent, waar de zorg integraal intramuraal aangeboden kan worden. Hoewel ze vaak voorkwamen, gingen deze ontvluchtingen slechts in een op vier van de gevallen gepaard met het plegen van een als misdrijf omschreven feit, in meerderheid niet gewelddadige feiten zoals drugbezit of diefstal.

Iets meer dan een derde van de incidenten betrof strafbare feiten (37.4%). Drie kwart van die feiten vond plaats op de MSA (74.8%) en een vierde vond plaats in de samenleving. Geïnterneerden pleegden significant meer feiten op de afdeling dan tijdens vrijheden in de samenleving (McNemar, p = .00). Ook het aandeel geweldsfeiten was hoger op de MSA dan daarbuiten (58.0% versus 30.0%). Het aandeel druggerelateerde feiten en eigendomsdelicten lag dan weer hoger buiten de MSA dan op de afdeling (respectievelijk 23.6% versus 40.0% en 9.7% versus 26.0%).

In navolging van internationaal onderzoek blijkt dat de meest voorkomende strafbare feiten geweldsincidenten waren, waarvan het merendeel verbaal geweld of slagen en verwondingen betrof (Blattner & Dolan, 2009; Gow et al., 2010). Bij alle incidenten was slechts drie keer sprake van zeer ernstig geweld, met name twee pogingen tot een levensdelict en een brandstichting met gevaar voor personen. Conform eerder onderzoek (Gudjonsson, Rabe-Hesketh, & Wilson, 1999), kan dus gesteld worden dat de aard van de geweldsincidenten in het merendeel van de gevallen een lage tot matige ernst kent. Het lage percentage fysieke geweldsincidenten is opvallend vermits de meerderheid van de onderzochte populatie in het verleden reeds werd veroordeeld en/of geïnterneerd voor geweldsdelicten. Deze vaststelling kan wijzen op een adequaat risicomanagement van de MSA's, waar men onder meer gebruik maakt van individuele signaleringsplannen en agressiehanteringstherapie om geweldsincidenten te voorkomen. Bij gewelddadige feiten was het slachtoffer bijna altijd een bekende van de dader. In de eerste plaats betrof het hulpverleners en in de tweede plaats medepatiënten, wat buitenlands onderzoek bevestigt (Gow et al., 2010; Gradillas et al., 2007).

Incidenten leiden tot vrijheidsberoving, niet tot een nieuwe veroordeling of internering

Bij iets meer dan de helft van de incidentmeldingen (54.3%) besliste de CBM tot een verderzetting van de MSA opname. Hoewel na een incident slechts uitzonderlijk een nieuwe strafrechtelijke veroordeling of interneringsmaatregel werd uitgesproken, volgde na een incident in iets minder dan de helft van de gevallen toch een scherpe strafrechtelijke reactie – zelfs zonder vonnis – in de vorm van een (her)opsluiting in de gevangenis (wederopname). Wederopnames werden, conform de wet, bevolen door de Procureur des Konings meestal in overleg met de CBM, maar altijd op vraag van de MSA's zelf (H. Heimans, persoonlijke communicatie, 10 december 2014). Opvallend is dat bij de beslissing tot een wederopname in de gevangenis geen significant verschil werd gevonden naargelang de aard van de feiten. Meldingen van strafbare incidenten en niet strafbare incidenten leidden in respectievelijk 52.0% en 39.0% van de gevallen tot een wederopname in de gevangenis, maar dit verschil was niet significant. Een wederopname na een incident had een langdurige detentie, gemiddeld genomen bijna 15 maanden, tot gevolg. De aangevatte behandeling werd hierdoor bij een derde van de populatie voor langere tijd onderbroken en bij een vierde van de populatie zelfs (voorlopig) stopgezet. Van de optie om rechtstreeks te kiezen voor een tweede behandelpoging in een van de andere MSA's werd nagenoeg geen gebruik gemaakt. Een langdurige investering geleverd tijdens een behandeltraject kan hierdoor op relatief korte termijn teniet gedaan worden door detentieschade (Crewe, 2011; Lauwaert, Mattheeuws, & De Deygere, 2014). Carr et al. (2006) verklaren het hoog aantal vastgelopen behandelingen door een gebrekkige voorbereiding tijdens detentieperiode op forensisch psychiatrische behandeling waardoor patiënten snel worden teruggestuurd naar de gevangenis. Bovendien kunnen patiënten die in detentie verbleven, specifiek gedrag vertonen (zoals wantrouwen tegenover het personeel en het verzwijgen van klachten), dat het therapeutisch proces kan bemoeilijken en de kans op incidenten kan vergroten. Deze problemen waren mede de aanleiding voor de opstart van de schakelteams internering die proberen op casusniveau zorgtrajecten uit te werken voor geïnterneerden bij wie het uitwerken van een zorgtraject moeilijk verloopt (Lauwaert et al., 2014). Een ander pilootproject vanuit de MSA's zelf, het *InReach* project, werd gestart om opnames vanuit de gevangenis te faciliteren en drop-out te voorkomen (Stassen, Habets, Mertens, De Laender, & Jeandarme, 2014). Een psychiatrisch verpleegkundige

vormt in dit project de brugfunctie tussen het ziekenhuis en de gevangenis en werkt pretherapeutisch en motivationeel.

Een tweede interessante bevinding is dat meldingen van geweldsincidenten slechts zelden (4.3%, *n* = 2) tot een nieuwe veroordeling/internering leidden. Dit bleek ook uit recent onderzoek in de Nederlandse GGZ (Harte et al., 2013). Een mogelijke verklaring die de auteurs hiervoor geven, is het vervolgingsbeleid van het Openbaar Ministerie (OM). Zo kan het OM op grond van het opportuniteitsbeginsel beslissen een strafbaar feit niet te vervolgen, bijvoorbeeld wanneer het minder ernstige feiten betreft. Uit ons onderzoek bleek inderdaad dat het merendeel van de strafbare feiten, ook de gewelddadige feiten, als minder ernstig kunnen beschouwd worden. Daarnaast staan geïnterneerden al onder een beschermingsstatuut, waardoor mogelijk wordt geopteerd voor een reactie via het huidige interneringssysteem, zoals een vermaning of een wederopname in de gevangenis, in plaats van een vervolging. Op deze manier wordt een snel en flexibel justitieel ingrijpen mogelijk gemaakt.

Geweldsdelicten hangen samen met een moeizaam behandelverloop

Wanneer de groep geïnterneerden die betrokken zijn bij geweldsdelicten van naderbij wordt bekeken, kunnen enkele relevante zaken vastgesteld worden. In vergelijking met de groep geïnterneerden die niet betrokken waren bij geweldsdelicten, kende de eerstgenoemde groep vaker een problematisch behandelverloop, gekenmerkt door moeilijkheden bij het nakomen van behandelafspraken en/of een vroegtijdige beëindiging van de behandeling. In eerder onderzoek bij psychiatrische patiënten werd reeds een verband vastgesteld tussen therapietrouw en het voorkomen van geweldsincidenten (Monahan et al., 2001; O'Farrell, Murphy, Stephan, Fals-Stewart, & Murphy, 2004). Voor zover bekend, werden nog geen studies uitgevoerd bij forensisch psychiatrische populaties met een specifieke focus op het verband tussen therapietrouw en het voorkomen van geweld. In de Nederlandse studie van Mudde et al. (2011) bleek wel dat het niet reageren op behandeling een verband aantoonde met agressief gedrag tijdens behandeling. Aangezien uit de resultaten bleek dat een gebrek aan compliance, een mogelijke voorspeller was van een vroegtijdig beëindigde behandeling en geassocieerd was met geweldsincidenten, valt het aan te bevelen om (nog) meer te investeren in het therapeutische klimaat en geïnterneerden beter voor te bereiden op behandeling. Cornaggia, Beghi, Pavone, and Barale (2011) benadrukten

in dit verband het belang van een warm behandelklimaat, gekenmerkt door een gepast aantal getrainde verpleegkundigen en geen overbevolking.

Naast een problematisch behandelverloop, vertoonde alcoholgebruik tijdens de behandeling een verband met geweldsincidenten. Alcoholgebruik is in principe verboden tijdens MSA behandeling; zo nodig worden specifieke therapieën voor middelenmisbruik ingezet. Ook in de literatuur wordt alcoholgebruik in verband gebracht met geweld van psychiatrische patiënten tijdens opname, hoewel andere factoren, zoals het voorkomen van geweldsdelicten in het verleden, een gedwongen opname en bepaalde symptomen zoals impulsiviteit en bevelshallucinaties, belangrijker geacht worden in het voorspellen van agressie (Cornaggia et al., 2011; Steinert, 2002).

De vaststelling dat medium security geïnterneerden die betrokken waren bij geweldsincidenten vaker een klinisch gestructureerd oordeel *hoog* op de HCR-20 hadden dan medium security geïnterneerden die niet betrokken waren bij geweldsincidenten blijft niet overeind in het meervoudig regressiemodel. Het ontbreken van een verband moet echter genuanceerd worden omdat bij ongeveer vier op de tien geïnterneerden geen HCR-20-score beschikbaar was. Bovendien werd in de literatuur meermaals een verband gevonden tussen een hoge score op de HCR-20 en het voorkomen van geweldsincidenten (McDermott et al., 2008; Mudde et al., 2011).

Geweldsincidenten bleken in deze studie niet samen te hangen met een aantal eerder beschreven risicofactoren voor herval zoals geslacht, leeftijd, geboorteland, type indexdelict en de omvang van het strafrechtelijk verleden (Wartna, el Harbachi, & van der Knaap, 2005). Een mogelijke verklaring is dat de meeste van deze risicofactoren herval na behandeling voorspellen, eerder dan tijdens behandeling (Wang & Diamond, 1999). Evenmin werd een verband gevonden met de aard van de psychopathologie. In navolging van recent onderzoek in Nederland was een diagnose middelenmisbruik/-afhankelijkheid niet geassocieerd met een verhoogd risico op geweldsincidenten (van der Kraan et al., 2014).

Er werd geen verband teruggevonden tussen geweldsincidenten en een hoge PCL-R score. Ook hier dient het ontbreken van een verband genuanceerd te worden omdat bij ongeveer zes op de tien geïnterneerden geen PCL-R score werd bepaald. Sommige studies hebben wel een associatie gevonden tussen fysiek gewelddadige incidenten en een hoge psychopathiescore;

andere studies geven aan dat dit verband niet eenduidig is (Guy et al., 2005). In een studie bij Nederlandse TBS-gestelden was een hoge PCL-R score niet gerelateerd aan fysieke agressie tijdens de opname, maar wel aan een moeizamer behandelverloop waarin meer incidenten en verbale agressie plaatsvonden (Hildebrand, de Ruiter, & Nijman, 2004).

Beperkingen van het onderzoek

Aan deze studie zijn een aantal beperkingen verbonden waardoor voorzichtigheid geboden is bij de interpretatie en bij veralgemening van de bevindingen. Ten eerste werd uitgegaan van officieel geregistreerde incidenten die door de MSA's gemeld werden aan de justitie-assistent en/of de CBM. Het is dan ook onduidelijk of en hoe groot het dark number is. Er bestaat in België vooralsnog geen aangifteplicht van incidenten door forensisch psychiatrische voorzieningen. Toch is het zeer onwaarschijnlijk dat justitieassistenten of de CBM, gelet op hun controlefunctie, niet door de MSA's op de hoogte zouden gesteld worden van de meer ernstige incidenten. Evenmin hebben de verschillende MSA's een uniforme regeling omtrent de registratie van incidenten. Daarnaast is het onduidelijk of hulpverleners in de Vlaamse forensische GGZ tolerant staan tegenover grensoverschrijdend gedrag met een onderrapportage van voornamelijk verbale geweldsincidenten als gevolg (Gow et al., 2010).

Ten tweede was geen controlegroep voorhanden zodat geen vergelijkingen konden worden gemaakt met patiënten in de algemene en/of de andere vormen van forensische GGZ. In Vlaanderen vonden slechts twee studies plaats rond incidenten binnen de forensische GGZ. Omwille van hun specifieke onderzoekspopulatie (geïnterneerden met een verstandelijke beperking en jongeren) en prospectief onderzoeksdesign is een vergelijking met dit onderzoek niet aan de orde (Pouls & Jeandarme, 2013; Tremmery, de Decker, De Hert, De Varé, & Danckaerts, 2012). Hoewel getracht werd de bevindingen te relateren aan internationale studies bij medium security populaties, is ook hier de nodige omzichtigheid geboden wegens verschillen in methodologie (zoals een bredere operationalisering van het begrip *incident* in de huidige studie) en in kenmerken van de onderzoekspopulaties.

Ten derde is het onderzoek retrospectief en werd gebruik gemaakt van bestaande dossiers. De kwaliteit van dossiers is niet steeds optimaal omdat deze vertekeningen kunnen bevatten en enkel die gegevens die daadwerkelijk werden geregistreerd kunnen gebruikt worden

(Lievens, 2001). Tijdens de dataverzameling bleek dat bepaalde informatie niet systematisch beschikbaar was. Dit was met name het geval voor de risicotaxatie-instrumenten en IQ gegevens, die in de Vlaamse praktijk niet standaard worden afgenomen. Bovendien geeft een dossierstudie geen inzicht in de manier waarop incidenten ervaren worden door patiënten en hulpverleners en wat de invloed is van deze incidenten op de relatie patiënt-hulpverlener (Cornaggia et al., 2011).

Ten vierde is waarschijnlijk sprake van een populatie bias. Geïnterneerden met een primaire seksuele, verslavings- of psychopathische problematiek kwamen immers in principe niet in aanmerking voor MSA opname. Dit kan misschien de lage prevalentie van seksuele delicten bij incidenten verklaren.

Tot slot werd ervoor gekozen om de data voor de drie instellingen samen te publiceren. Hierdoor kon de invloed van mogelijke verschillen in behandelvisie en -klimaat niet weergegeven worden.

CONCLUSIE

Met de opstart van de medium security afdelingen werd in Vlaanderen eindelijk werk gemaakt van een aanzet tot een forensisch psychiatrisch circuit. Boers, Vandevelde, Soyez, De Smet, and To (2011) schetsten een beeld van de aangeboden behandelingen, maar onderzoek naar medium security geïnterneerden of hun behandelingsverloop was tot nog toe afwezig. De voorliggende studie onderzocht voor de eerste keer met welke incidenten deze afdelingen de eerste tien jaren van hun werking geconfronteerd werden. Bijzondere aandacht werd hierbij besteed aan geweldsincidenten omdat zij een voorziening in het bijzonder kunnen destabiliseren en zorgen voor een groot personeelsverloop.

Ondanks de beperkingen van het onderzoek kunnen enkele conclusies getrokken worden. In de eerste plaats bleek dat, ondanks de negatieve beeldvorming over forensisch psychiatrische patiënten (Gradillas et al., 2007; Lamb et al., 1999), bij bijna de helft (46.8%) van de patiënten tijdens een gemiddelde behandelperiode van 1.7 jaar geen enkel incident werd gemeld aan de CBM. Daarenboven betrof het merendeel van de gerapporteerde incidenten geen strafbare (62.6%) noch fysiek agressieve feiten. We kunnen dus stellen dat relatief veilig werken met deze groep patiënten mogelijk is.

Het risicomanagement van de MSA's bleek afdoende om de grootste groep geïnterneerden, ondanks hun complexe problematiek, te behandelen. Voor de MSA teams is dit een bevestiging dat hun inspanningen lonen. Bij een subpopulatie werd de behandeling echter onderbroken en/of stopgezet en werden geïnterneerden terug gedetineerden. Deze drop-out geïnterneerden vormen een groep die in de toekomst nader moet onderzocht worden. De komst van het forensisch psychiatrisch centrum (FPC) in Gent zou in de toekomst voor deze geïnterneerden een oplossing kunnen bieden. Dit vraagt wel dat voorzien wordt in een wettelijke mogelijkheid tot wederopsluiting vanuit een MSA in een FPC in plaats van in een gevangenis. Dat hier niet voor gekozen werd, is ons inziens een gemiste kans. De medium security afdelingen zelf namen reeds initiatieven zoals het InReach project. Door meer te investeren in de therapeutische relatie en patiënten beter voor te bereiden op behandeling, hoopt men deze vastgelopen behandelingen te voorkomen.

Tot slot moet er blijvende aandacht uitgaan naar de hulpverleners en medepatiënten die de meeste kans lopen om slachtoffer van een incident te worden. Tot nog toe is er bijvoorbeeld geen beleid binnen de MSA rond aangifte of vervolging bij ernstige incidenten. Ook worden (gewelds)incidenten niet op een gestandaardiseerde manier geregistreerd, hoewel dit risico-inschattingen kan verbeteren en geweldsincidenten kan voorkomen (Cornaggia et al., 2011).

NOTEN

1 Wet 1 juli 1964 tot Bescherming van de Maatschappij tegen abnormalen en de gewoontemisdadigers, BS 17 juli 1964. Op 9 juli 2014 werd in het Belgisch Staatsblad de wet van 5 mei 2014 betreffende de internering van personen gepubliceerd. Deze wet zal uiterlijk op 1 januari 2016 in werking treden en de wet van 1964 opheffen en vervangen.

2 Relationele beveiliging omvat zowel kwantitatieve (zoals bijvoorbeeld patiënt-personeelsratio) als kwalitatieve aspecten.

3 Voor geïnterneerden met seksuele problematiek (zoals bijvoorbeeld pedofilie) en een matig risicoprofiel is behandeling mogelijk in een van de drie afdelingen voor seksueel delinquenten (Asster Sint-Truiden, Sint-Amandus Beernem en Sint-Lucia Sint-Niklaas). In tegenstelling tot de medium security afdelingen zijn voornoemde afdelingen niet exclusief voor geïnterneerden bedoeld.

4 Alle Vlaamstalige CBM's en de drie Vlaamse MSA's zijn betrokken bij dit onderzoek (*N* = 531). De dataverzameling werd recent afgerond; de komende periode worden de resultaten van de totale populatie geanalyseerd.

5 In totaal werden in deze periode 206 personen behandeld waarvan er drie (1.46%) geen toestemming gaven tot het gebruik van hun gegevens.

6 Het indexdelict is het delict waarvoor de internering werd opgelegd.

7 Extractie van de data op 16/11/2011.

8 Extractie van de data op 16/02/2012.

9 De afbakening van de variabelen werd opgesteld na consensusbespreking door twee onderzoekers (tabel 1).

10 Aangezien het geslacht niet geassocieerd was met de onderzochte variabelen en omwille van het kleine aantal vrouwelijk geïnterneerden werden de onderzoeksresultaten niet apart weergegeven voor mannen en vrouwen.

11 De overgrote meerderheid (98.36%) hiervan had ook de Belgische nationaliteit.

12 De opnameduur werd per geïnterneerde voor eerste opname en heropnames gecumuleerd.

13 Een geïnterneerde werd als een veelpleger beschouwd als hij als volwassen dader drie of meer strafzaken onderging binnen vijf jaar voorafgaand aan de indexinternering (Wartna &

Tollenaar, 2004). Bij 17.2% van de populatie (*n* = 35) was omwille van de jonge leeftijd geen periode van vijf jaar of meer voor de indexinternering.

[14] Deze indeling is gebaseerd op categorie 8 – 12 (Brand, 2005).

[15] Categorie a) en b) al dan niet in combinatie met een persoonlijkheidsstoornis.

[16] De HCR-20 is een lijst van 20 risicofactoren voor gewelddadig gedrag. De HCR-20 is geen formele psychologische test, maar een handleiding voor risicotaxatie. Op basis van 20 variabelen die relevante aspecten van zowel verleden, heden als toekomst beschrijven, wordt een inschatting gemaakt van de kansen op herval in een geweldsdelict (laag/matig/hoog). Een gestructureerd klinische score *hoog* op de HCR-20 geeft een indicatie op een verhoogd risico op herval in geweldsincidenten.

[17] De hoogte van deze score geeft een indicatie van de mate waarin psychopathische kenmerken aanwezig zijn. De maximale score op de PCL-R bedraagt 40. Een score van 25 of meer wordt in Europa als indicatief voor psychopathie aanzien (Cooke & Michie, 1999).

[18] De beschrijvende statistiek en de analyses werden uitgevoerd in SAS 9.3 en R 3.0.1. De continue variabelen werden beschreven door middel van het gemiddelde, de standaarddeviatie en de range. De categorische variabelen werden beschreven door middel van de aantallen en percentages per categorie. De associatie tussen twee categorische variabelen werd getest met de Fisher-exact test. De gemiddelden van twee groepen (eventueel na transformatie van de data in het geval van niet-normaliteit) werden vergeleken met de ongepaarde t-test. Indien geen gepaste transformatie van de data gevonden werd bij niet-normaliteit, werden de medianen van de twee groepen vergeleken met de Wilcoxon rank-sum test. De associaties tussen geweldsincidenten tijdens MSA behandeling en de verschillende justitiële, klinische en behandelingskenmerken werden eerst individueel getest met de Fisher-exact test. Het bijbehorende relatieve risico (RR) met het 95% betrouwbaarheidsinterval werd berekend. Vervolgens werd, op basis van de gevonden enkelvoudige associaties, de associatie tussen incidenten tijdens de MSA behandeling en de verschillende justitiële, klinische en behandelingskenmerken onderzocht met behulp van een meervoudige logistische regressie. De relatieve risico's voor het meervoudige model werden berekend op basis van de gecorrigeerde odds ratio's volgens de methode beschreven door (Zhang & Yu, 1998). Alle testen waren tweezijdig met significantieniveau 5%.

[19] Van twee van de in totaal 331 incidenten (0.6%) kon geen juiste kwalificatie achterhaald worden.

[20] Van 119 ontvluchtingen was de begin- en einddatum bekend. Daarnaast was bij een ontvluchting de geïnterneerde bij het afsluiten van de studie na 2833 dagen nog steeds spoorloos.

[21] In slechts 39.1% van de gevallen (*n* = 25) was het geslacht van het slachtoffer bekend. In vier gevallen (6.3%) kon de aard van de relatie tussen dader en slachtoffer niet achterhaald worden.

[22] Het ging hier over hulpverleners in de ruime zin: verpleegkundigen, therapeuten, sociaal assistenten, bewindvoerders, …

[23] In twee van de 236 meldingen (0.9%) kon de precieze kwalificatie van het meest ernstige incident per melding niet achterhaald worden.

[24] In deze gevallen kon de geïnterneerde wel vermaand worden op de CBM-zitting.

[25] De detentieperiode was bij 13 geïnterneerden die werden heraangehouden niet beëindigd op 31/12/2010. De gemiddelde detentietijd werd berekend tot en met 31/12/2010.

REFERENTIES

Abidin, Z., Davoren, M., Naughton, L., Gibbons, O., Nulty, A., & Kennedy, H. G. (2013). Susceptibility (risk and protective) factors for in-patient violence and self-harm: Prospective study of structured professional judgement instruments START and SAPROF, DUNDRUM-3 and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry, 13*, 197. doi: 10.1186/1471-244x-13-197

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Blattner, R., & Dolan, M. (2009). Outcome of high security patients admitted to a medium secure unit: The Edenfield Centre study. *Medicine, Science, and the Law, 49*(4), 247-256. doi: 10.1258/rsmmsl.49.4.247

Boers, A., Vandevelde, S., Soyez, V., Smet, S. De, & To, W. T. (2011). Het zorgaanbod voor geïnterneerden in België. *Panopticon, 32*(2), 17-38.

Brand, E. F. J. M. (2005). *Onderzoeksrapport PIJ-dossier 2003-C. Predictieve validiteit van de FPJ-lijst*. Den Haag: DJI.

Carr, W. A., Rotter, M., Steinbacher, M., Green, D., Dole, T., Garcia-Mansilla, A., . . . Rosenfeld, B. (2006). Structured Assessment of Correctional Adaptation (SACA): A measure of the impact of incarceration on the mentally ill in a therapeutic setting. *International Journal of Offender Therapy and Comparative Criminology, 50*(5), 570-581. doi: 10.1177/0306624X06289176

Cooke, D. J., & Michie, C. (1999). Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology, 108*(1), 58–68. doi: 10.1037/0021-843X.108.1.58

Cornaggia, C. M., Beghi, M., Pavone, F., & Barale, F. (2011). Aggression in psychiatry wards: A systematic review. *Psychiatry Research, 189*(1), 10-20. doi: 10.1016/j.psychres.2010.12.024

Cosyns, P., D'Hont, C., Janssens, D., Maes, E., & Verellen, R. (2007). Geïnterneerden in België. De cijfers. *Panopticon, 28*(1), 46-61.

Cosyns, P., Dillen, C., Ruyter, B. De, & Doncker, D. De (1994). Behandelingsmogelijkheden van geïnterneerden in algemeen psychiatrische ziekenhuizen. In F. Lampaert (Ed.), *Gevangenis en samenleving II* (pp. 85-91). Brussel: Koning Boudewijnstichting.

Crewe, B. (2011). Depth, weight, tightness: Revisiting the pains of imprisonment. *Punishment & Society, 13*(5), 509-529. doi: 10.1177/1462474511422172

Gadon, L., Johnstone, L., & Cooke, D. (2006). Situational variables and institutional violence: A systematic review of the literature. *Clinical Psychology Review, 26*(5), 515-534. doi: 10.1016/j.cpr.2006.02.002

Gow, R. L., Choo, M., Darjee, R., Gould, S., & Steele, J. (2010). A demographic study of the Orchard Clinic: Scotland's first medium secure unit. *The Journal of Forensic Psychiatry & Psychology, 21*(1), 139-155. doi: 10.1080/14789940903188956

Gradillas, V., Williams, A., Walsch, E., & Fahy, T. (2007). Do forensic psychiatric inpatient units pose a risk to local communities. *The Journal of Forensic Psychiatry & Psychology, 18*(2), 261-265. doi: 10.1080/14789940701246360

Gudjonsson, G., Rabe-Hesketh, S., & Wilson, C. (1999). Violent incidents on a medium secure unit over a 17-year period. *The Journal of Forensic Psychiatry, 10*(2), 249-263. doi: 10.1080/09585189908403679

Guy, L. S., Edens, J. F., Anthony, C., & Douglas, K. S. (2005). Does psychopathy predict institutional misconduct among adults? A meta-analytic investigation. *Journal of Consulting and Clinical Psychology, 73*(6), 1056-1064. doi: 10.1037/0022-006X.73.6.1056

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist*. Toronto, ON: Multi-Health Systems.

Harte, J. M., Leeuwen, M. E. van, & Theuws, R. (2013). Agressie en geweld tegen hulpverleners in de psychiatrie; aard, omvang en strafrechtelijke reactie [The nature, extent and judicial response to aggression and violence directed against care workers in psychiatry]. *Tijdschrift voor Psychiatrie, 55*(5), 325-335.

Hildebrand, M., Ruiter, C. de, & Nijman, H. (2004). PCL-R psychopathy predicts disruptive behavior among male offenders in a Dutch forensic psychiatric hospital. *Journal of Interpersonal Violence, 19*(1), 13-29. doi: 10.1177/0886260503259047

Hillbrand, M. (1995). Aggression against self and aggression against others in violent psychiatric patients. *Journal of Consulting and Clininal Psychology, 63*(4), 668-671. doi: 10.1037/0022-006X.63.4.668

Kennedy, H. G. (2002). Therapeutic uses of security: Mapping forensic mental health services by stratifying risk. *Advances in Psychiatric Treatment, 8*, 433-443. doi: 10.1192/apt.8.6.433

Kraan, J. van der, Verkes, R. J., Goethals, K., Vissers, A., Brazil, I., & Bulten, E. (2014). Substance use disorders in forensic psychiatric patients. *International Journal of Law and Psychiatry, 37*(2), 155-159. doi: 10.1016/j.ijlp.2013.11.009

Laenen, F. Vander, & Cauwer, B. De (2011). Just-care, justitie en zorg in dialoog over internering. *Panopticon, 32*(3), 55-59.

Lamb, H. R., Weinberger, L. E., & Gross, B. H. (1999). Community treatment of severely mentally ill offenders under the jurisdiction of the criminal justice system: A review. *Psychiatric Services, 50*(7), 907-913.

Lauwaert, F., Mattheeuws, K., & Deygere, M. De (2014). Het schakelteam internering: Naar een completer en dynamisch zorglandschap voor geïnterneerden. *Panopticon, 35*(3), 237-241.

Lievens, J. (2001). Werken met secundaire data. In J. Billiet & H. Waege (Eds.), *Een samenleving onderzocht: Methoden van sociaal-wetenschappelijk onderzoek* (pp. 343-363). Antwerpen: Standaard Uitgeverij.

McDermott, B. E., Edens, J. F., Quanbeck, C. D., Busse, D., & Scott, C. L. (2008). Examining the role of static and dynamic risk factors in the prediction of inpatient violence: Variable- and person-focused analyses. *Law and Human Behavior, 32*(4), 325-338. doi: 10.1007/s10979-007-9094-8

Moens, I., & Pauwelyn, L. (2012). *Geen opsluiting, maar sleutels tot re-integratie. Voorstellen voor een gecoördineerd zorgtraject voor geïnterneerden*. Brussel: Zorgnet Vlaanderen.

Monahan, J., Steadman, H. J., Silver, E., Appelbaum, P., Robbins, P., Mulvey, E., . . . Banks, S. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York, NY: Oxford University Press.

Mudde, N., Nijman, H., Hulst, W. van der, & Bout, J. van den (2011). Het voorspellen van agressie tijdens de behandeling van forensisch psychiatrische patienten aan de hand van de HCR-20 [Predicting aggression during the treatment of forensic psychiatric patients by means of the HCR-20]. *Tijdschrift voor Psychiatrie, 53*(10), 705-713.

O'Farrell, T. J., Murphy, C. M., Stephan, S. H., Fals-Stewart, W., & Murphy, M. (2004). Partner violence before and after couples-based alcoholism treatment for male alcoholic patients: The role of treatment involvement and abstinence. *Journal of Consulting and Clinical Psychology, 72*(2), 202-217.

Pouls, C., & Jeandarme, I. (2013). Institutional aggression in intellectually disabled (ID) offenders - An analysis of 5 specialized forensic ID projects. In N. O. P. Callaghan, J. H BjØrngaard, H. Nijman, T. Palmstierna, R. Almvik, & B. Thomas (Ed.), *Proceedings of the 8th European Congress on Violence in Clinical Psychiatry.* (pp. 408-411). Amsterdam: Kavanah.

Schalock, R. L., Borthwick-Duffy, S. A., Bradley, V. J., Buntinx, W. H. E., Coulter, D. L., Craig, E. M., . . . Yeager, M. H. (2010). *Intellectual disability: Definition, classification, and systems of supports*. Washington, DC: American Association on Intellectual and Developmental Disabilities.

Schanda, H., Stompe, T., & Ortwein-Swoboda, G. (2009). Dangerous or merely difficult? The new population of forensic mental hospitals. *European Psychiatry, 24*(6), 365-372. doi: 10.1016/j.eurpsy.2009.07.006

Smedt, A. De, Mariën, L., & Vermeiren, R. (2008, September). *Profilering van de medium risk geïnterneerden in forensische psychiatrische settings.* Paper presented at the Fourth Vlaams Geestelijk Gezondheidscongres, Ghent, Belgium.

Stassen, W., Habets, P., Mertens, A., Laender, J. De, & Jeandarme, I. (2014). The InReach project: From penitentiary to forensic hospital. *Therapeutic Communities: The International Journal of Therapeutic Communities 35*(3), 119-126.

Steinert, T. (2002). Prediction of inpatient violence. *Acta Psychiatrica Scandinavica.*(412), 133-141.

Tremmery, S., Decker, A. de, Hert, M. De, Varé, J. De, & Danckaerts, M. (2012). Agressiebeleid op een forensische psychiatrische unit voor adolescenten; Literatuuroverzicht en implementatie in de

praktijk [Managing aggression in a forensic psychiatric unit for adolescents: Literature review and clinical implementation]. *Tijdschrift voor Psychiatrie, 54*(9), 797-806.

Vandevelde, S., Soyez, V., Beken, T. Vander, Smet, S. De, Boers, A., & Broekaert, E. (2011). Mentally ill offenders in prison: The Belgian case. *International Journal of Law and Psychiatry, 34*(1), 71-78. doi: 10.1016/j.ijlp.2010.11.011

Wang, E. W., & Diamond, P. M. (1999). Empirically identifying factors related to violence risk in corrections. *Behavioral Sciences & the Law, 17*(3), 377-389.

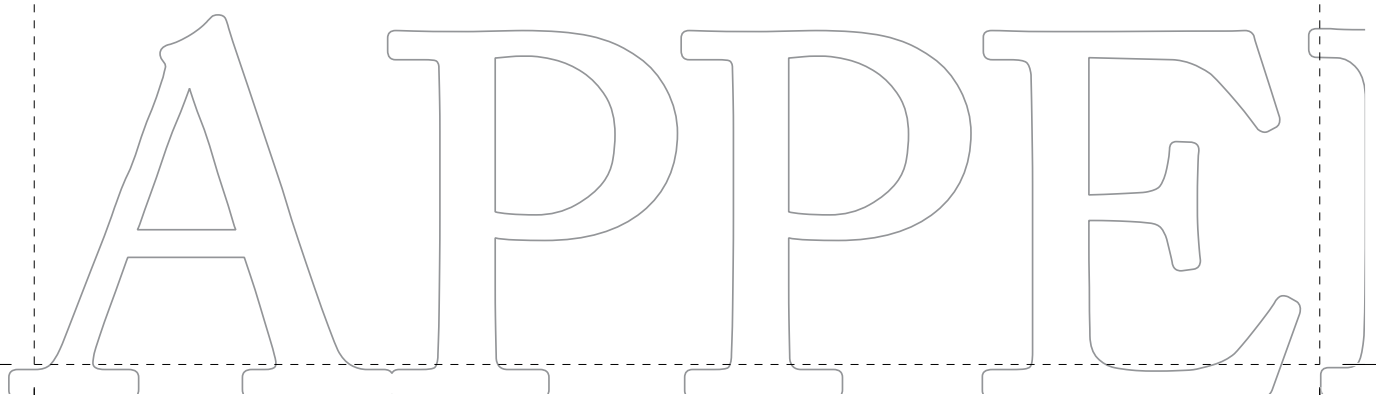
Wartna, B. S. J., Harbachi, S. el, & Knaap, L. M. van der (2005). *Buiten behandeling. Een cijfermatig overzicht van de strafrechtelijke recidive van ex-terbeschikkinggestelden*. Den Haag: Boom Juridische uitgevers.

Wartna, B. S. J., & Tollenaar, N. (2004). *Bekenden van Justitie. Een verkennend onderzoek naar de 'veelplegers' in de populatie van vervolgde daders*. Den Haag: Boom Juridische uitgevers.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence* (Version 2). Burnaby, BC: Simon Fraser University.

Wechsler, D. (2000). *WAIS-III Nederlandstalige bewerking. Afname en scoringshandleiding*. Lisse: Swets & Zeitlinger.

Zhang, J., & Yu, K. F. (1998). What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *Journal of the American Medical Association, 280*(19), 1690-1691. doi: 10.1001/jama.280.19.1690



Intelligence is in the eye of the beholder: Investigating repeated IQ measurements in forensic psychiatry

Habets, P., Jeandarme, I., Uzieblo, K., Oei, K., & Bogaerts, S.

Journal of Applied Research in Intellectual Disabilities, 28(3), 182-192

APPENDIX

ABSTRACT

Background. A stable assessment of cognition is of paramount importance for forensic psychiatric patients (FPP). The purpose of this study was to compare repeated measures of IQ scores in FPPs with and without intellectual disability.

Methods. Repeated measurements of IQ scores in FPPs (*n* = 176) were collected. Differences between tests were computed, and each IQ score was categorized. Additionally, t-tests and regression analyses were performed.

Results. Differences of 10 points or more were found in 66% of the cases comparing WAIS-III with RAVEN scores. Fisher's exact test revealed differences between two WAIS-III scores and the WAIS categories. The WAIS-III did not predict other IQs (WAIS or RAVEN) in participants with intellectual disability.

Discussion. This study showed that stability or interchangeability of scores is lacking, especially in individuals with intellectual disability. Caution in interpreting IQ scores is therefore recommended, and the use of the unitary concept of IQ should be discouraged.

INTRODUCTION

A stable assessment of cognitive functioning (i.e., intelligence) is of paramount importance for forensic psychiatric patients. The level of an individual's intellect has an impact on interrogations, court proceedings, court rulings, risk assessments and treatment programs. Consequently, countries have specific procedures regarding offenders with intellectual disability (OIDs). For example, in Belgium and the Netherlands, if an individual who committed a crime has a diagnosis of intellectual disability, it is possible that he/she will not be held responsible for his/her actions (not guilty by reason of insanity). As a result, a protection measure will be ordered (van Emmerik, 2001; Verlinden, Maes, & Goethals, 2009). Additionally, in most states of the United States, people with OIDs are not allowed to be executed. The assessment of intellectual disability can therefore literally be a matter of life and death, leaving no room for error (Fabian, Thompson, & Lazarus, 2011). Despite these concerns, uniformity is still lacking in assessment of intelligence in forensic populations. Different tools that do not – or only partly – measure the same aspects of intelligence are used, resulting in poorly interchangeable scores (McBrien, 2003; Uzieblo, Winter, Vanderfaeillie, Rossi, & Magez, 2012). It is therefore critical that intelligence is measured in a valid and stable manner and composite scores should be avoided. Namely, it is widely acknowledged that intelligence has a hierarchical structure (i.e., the Cattell–Horn–Carroll model) (McGrew, 2009), and minimizing intelligence into a single score fails to captivate the complexity of a person's intellect especially in persons with borderline intelligence (Uzieblo et al., 2012). Although large correlations between IQ tests have been reported, research has shown that scores obtained on intelligence tests given to the same individual are not identical (Di Nuovo, Di Nuovo, & Buono, 2012; Floyd, Clark, & Shadish, 2008). In fact, IQ scores are not expected to have perfect instrumental or temporal stability (Evans, 1991). Studies regarding stability and consistency between and within IQ tests have shown positive results. For example, Wechsler (1997) reported a 0.91 stability coefficient (i.e., the correlation between assessments using the same test within the same individual) of the Wechsler Adult Intelligence Scale III (WAIS-III) with one month separating the two assessments (Wechsler, 1997). However, these coefficients were acquired in individuals within the normal IQ range within a short time period, making it not necessarily representative for individuals with intellectual disability. A meta-analysis by Whitaker (2008b)

investigated stability coefficients in individuals with intellectual disability and found reasonable stability for full-scale IQs (0.82). Despite the relative stability of scores, a 10-point change or more between assessments with the same instrument was found in 14% of the subjects. Investigating differences in IQ between instruments in an intellectual disability population, Silverman et al. (2010) found a mean difference between the WAIS (Wechsler, 1955) and Stanford-Binet (Roid, 2003) scores of 16.7 points in which the WAIS scored systematically higher than Stanford-Binet. A difference of 10 points or more was found in 85% of the individuals when comparing tests, and 24% had a 20-point difference or more. In contrast, they reported a strong correlation between the two tests (r = 0.82) indicating that, despite the large differences between the two instruments, they measured the same basic construct (Silverman et al., 2010). Research investigating stability of IQ scores within and between instruments in intellectual disability is scarce and even more so in forensic psychiatric populations. A recent Dutch study investigated the stability of IQ scores in a forensic psychiatric sample. IQ measurements – WAIS-III, Groninger Intelligence Test (GIT; Kooreman & Luteijn, 1987) and Kaufman Adolescent and Adult Intelligence Test (KAIT; Kaufman & Kaufman, 1993) – were collected and compared for 50 individuals. They found that when using the WAIS-III to determine intellectual disability, only eight individuals fell into the intellectual disability category, whereas when using the GIT and KAIT, 17 and 29 individuals, respectively, fell into the intellectual disability category. Additionally, about half of the individuals had a difference of at least 10 points when comparing the KAIT with the WAIS-III and the KAIT with the GIT (Van Toorn & Bon, 2011).

In sum, research has shown reasonable stability coefficients within tests and relatively high correlations between tests. However, large differences in IQ scores within individuals are possible, which consequently can have severe implications. Furthermore, the question remains whether studies of stability and interchangeability of IQ scores can be translated to the intellectual disability population. Recent evidence suggests that this might not be the case. The purpose of this study was to describe and compare repeated measurements of intelligence in a forensic psychiatric sample with and without intellectual disability. It was predicted that different IQ tests would result in different classifications of intellectual disability. Consequently, a different pattern of regression coefficients was expected to be found in individuals with intellectual disability when compared to individuals without intellectual disability.

METHOD

Sample and participant selection

This study is part of a large observational study, which is the first study in Flanders investigating recidivism in forensic psychiatric patients. Patients who were admitted between 2001 and 2010 to one of the three medium security forensic wards in Bierbeek, Rekem or Zelzate (n = 542) were eligible to be included in the study. Eleven patients refused participation, resulting in a final sample of 531 participants. Data were gathered by accessing prison and psychiatric hospital records. Information regarding level of education, psychiatric diagnosis, criminal history, hospitalization/imprisonment periods and IQ scores was collected. Diagnosis was based on the Diagnostic and Statistical Manual of mental disorders-IV-Text Revision (DSM-IV-TR; American Psychiatric Association, 2000).

Assessments and measures

The following intelligence tests were found: the Dutch adaptation of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955; Wechsler, 1970), the WAIS-III (Wechsler, 1997, 2005), Raven's Progressive Matrices (RAVEN; (Raven, Raven, & Court, 1998) and the short Groninger Intelligence Test (sGIT; Kooreman & Luteijn, 1987). Of the 531 participants, 176 (33%) had two or more IQ scores. The place of administration of the IQ tests is presented in Table 1. Reports of IQ tests can come from psychiatric centers, penitentiaries, and from forensic psychiatric assessments (FPAs). A FPA is ordered by a judge to assess whether or not the offender is accountable for his or her crimes and can include results of IQ tests. This assessment can take place in the penitentiary, in a psychiatric center or in an ambulatory setting. The mean age at the time of administration of the IQ test and the corresponding sample size are reported in Table 2.

WAIS and WAIS-III

The WAIS measures general intelligence or *g* and is divided into two parts: the verbal scale and the performance scale. Each of these two parts is further divided into subtests, each of which taps a specific verbal or non-verbal skill (Wechsler, 1955). The WAIS-III is the revised version of the WAIS-R (the successor of the WAIS). Given that the WAIS-R has never been translated into Dutch,

no scores are available for the WAIS-R. Two different Belgian norms from 2000 to 2005 are available for the WAIS-III (Tellegen, 2002; Wechsler, 2000; Wechsler, 2005). However, the norm table that was used was not found in the most recent majority of reports, making it impossible to recalculate the full-scale WAIS-III scores using norms.

Table 1. Number of IQ tests stratified for place of IQ test administration

	WAIS-III		WAIS		RAVEN		sGIT	
	# Obs.	%	# Obs.	%	# Obs.	%	# Obs.	%
Psychiatric center	126	75	30	49	9	12	0	
Penitentiary	12	7	11	18	3	4	0	
FPA	3	2	12	20	60	78	31	97
Other	2	1	4	7	0	0	0	
Unknown	24	14	4	7	5	6	1	3
Total	167		61		77		32	

Note. Only participants with more than one IQ score on record are included in this table. FPA = Forensic psychiatric assessment; # Obs. = number of observations per test (not equal to number of subjects as subjects have more than one IQ test).

Table 2. Differences in age at time of testing

	Mean (*SD*)	Mean difference (*SD*)	*t*	*df*	*n*
WAIS-III[(1)]	35.02 (9.57)	-1.42 (3.38)	-2.75	42	43
WAIS-III[(2)]	36.44 (9.00)				
WAIS-III	40.24 (8.86)	8.63 (5.43)*	11.69	53	54
WAIS	31.61 (9.43)				
WAIS-III	34.86 (8.17)	3.68 (3.99)*	7.75	70	71
RAVEN	31.18 (8.87)				
WAIS-III	31.41 (8.33)	1.34 (2.54)*	2.85	28	29
sGIT	30.07 (9.50)				

* $p < 0.01$.

RAVEN

The RAVEN is a nonverbal intelligence test that requires inductive reasoning about perceptual patterns and is considered to be a good measure for g and more specifically, *fluid g* (Schroth, 1983; Tulkin & Newbrough, 1968). Moreover, it has been shown to be a valid instrument in cross-cultural research (Jensen, 1980; Raven, Court, & Raven, 1983). Given that in Belgium many different norms are available (Moenaert, 2006), RAVEN raw scores were transformed using the latest Belgian norms (Magez, Moenaert, & Degezelle, 2006).

sGIT

The short version of the GIT2 (sGIT; Luteijn & Barelds, 2004) consists of six subtests (the full version contains 10 subtests) and is comparable to the WAIS. Studies have found a correlation of r = .94 between the sGIT and the GIT2, concluding that the sGIT can be administered without problems.

Statistical analyses

The WAIS-III was used as the reference score because it was the most frequently available score among participants who had more than one IQ score. When two WAIS-III scores are reported, the lowest score found in the database will be presented as WAIS-III$^{(1)}$ and the other WAIS-III score as WAIS-III$^{(2)}$. Difference scores were computed by subtracting the corresponding second IQ score from the WAIS-III score within a subject (WAIS-III$^{(1)}$ – WAIS-III$^{(2)}$; WAIS-III – WAIS; WAIS-III – RAVEN; WAIS-III – sGIT). Frequencies of the absolute difference scores are presented in Figure 1. Paired sample t-tests were performed to investigate whether WAIS-III$^{(1)}$ scores significantly differed from WAIS-III$^{(2)}$, WAIS, RAVEN or sGIT scores. For each IQ test, IQ scores were divided into categories: 1 = normal IQ (≥85), 2 = borderline IQ (71–84) and 3 = intellectual disability (≤70). Categorical differences between IQ scores were tested using Fisher's exact test. To investigate whether one IQ score predicted another IQ score, Figure 1 multilevel regression analyses were conducted using the XTREG command in STATA (StataCorp, 2011) because of the two-level grouping structure of the data, compromising statistical independence of the observations, namely IQ scores (level 1) were nested in subjects (level 2). WAIS-III$^{(1)}$ score was used as the independent variable, WAIS-III$^{(2)}$ score, WAIS score, RAVEN score, or sGIT score as the dependent variable and

subject number was modelled as random effect. Multilevel regression analyses were repeated with the year of IQ test administration as a covariate. To examine whether the level of association between IQ scores differed by education level or diagnosis of intellectual disability, multilevel regression analyses were repeated stratified by education level (1= normal education, 2= special needs education) and diagnosis of intellectual disability (1 = no diagnosis of intellectual disability, 2 = diagnosis of intellectual disability).

RESULTS

Demographic characteristics

Of the 167 participants, 5 (3%) were female. Participants exhibited the following Axis I diagnoses: developmental disorders (6%, *n* = 13), substance-related disorders (46%, *n* = 103), psychotic disorders (18%, *n* = 41), mood disorders (6%, *n* = 6), panic disorders (1%, *n* = 2), paraphilia (5%, *n* = 11), cognitive disorders (1%, *n* = 1), other disorders (17%, *n* = 37) and no or postponed diagnosis (4%, *n* = 9). Axis II diagnoses established in the participants were cluster A personality disorders (7%, *n* = 14), cluster B personality disorders (45%, *n* = 94), cluster C personality disorders (5%, *n* = 10), personality disorders not otherwise specified (13%, *n* = 28), intellectual disability (21%, *n* = 44) and no or postponed diagnosis (10%, *n* = 21). In total, 44 participants had a diagnosis of intellectual disability and 43 participants had been enrolled in special needs education (and four participants had missing values for education). Age at the time of testing differed significantly for all comparisons such that WAIS-III scores were from older individuals compared to the other tests (Table 2). In a number of cases, the amount of time between assessments was less than one year (WAIS-III$^{(1)}$ – WAIS-III$^{(2)}$ = 26%, WAIS-III – RAVEN = 25%, WAIS-III – sGIT = 24%).

Difference scores

When comparing the two WAIS-III scores, 33% of the cases had difference scores that were higher than 10 points. For sGIT scores, 60%, 66% and 52% of the cases, respectively, had difference scores higher than 10 points (Figure 1).

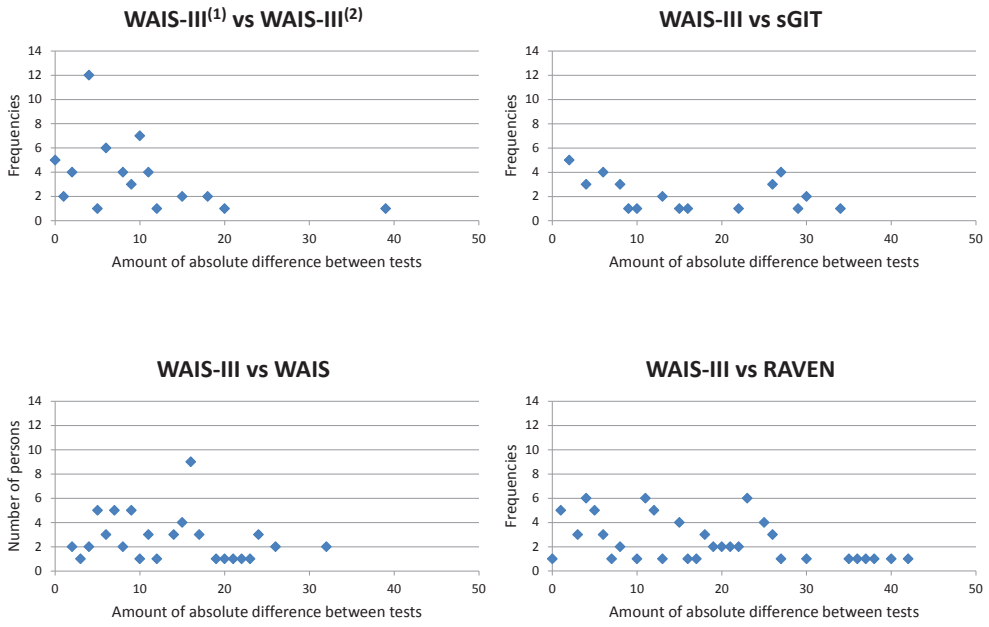


Figure 1. Frequencies of difference scores. Frequencies of absolute differences between tests are reported on the y-axis, and on the x-axis, the amount of absolute difference between tests is reported. For example, for the WAIS-III[(1)] versus WAIS-III[(2)], one person has a difference of 39 points, whereas 12 persons have a difference of four points and five persons have no differences between scores.

Differences between scores

All comparisons testing differences between IQ scores were significant (Table 3). The largest mean difference was found between the WAIS-III full-scale score and the sGIT (13.49), and the smallest mean difference was between the verbal IQ (VIQ) score on the WAIS-III[(1)] and the VIQ score on the WAIS-III[(2)] (5.25).

Change in category

Cross tabulation analyses using Fisher's exact test revealed significant differences in IQ categories between the WAIS-III[(1)] and the WAIS-III[(2)] and the WAIS-III and WAIS (Table 4). When comparing the WAIS-III[(1)] categories with the WAIS-III[(2)] categories, 15 of the 55 (27%) cases changed category: five (9%) from borderline to normal, nine (16%) from intellectual disability to borderline and one

Table 3. Differences and correlations between IQ scores

	Mean (*SD*)	Mean difference (*SD*)	*t*	*df*	*n*
WAIS-III[(1)] FSIQ	75.00 (14.67)	-7.44 (6.46)*	-8.53	54	0.90*
WAIS-III[(2)] FSIQ	82.44 (14.15)				
WAIS-III[(1)] VIQ	74.60 (17.06)	-5.25 (5.00)*	-6.88	42	0.96*
WAIS-III[(2)] VIQ	79.86 (16.94)				
WAIS-III[(1)] PIQ	75.39 (13.27)	-7.21(6.22)*	-7.6	42	0.89*
WAIS-III[(2)] PIQ	82.60 (13.33)				
WAIS-III FSIQ	83.36 (17.01)	-8.54 (12.44)*	-5.38	60	0.74*
WAIS FSIQ	91.90 (17.10)				
WAIS-III VIQ	81.56 (14.21)	-8.83 (8.49)*	-7.49	51	0.84*
WAIS VIQ	90.38 (15.58)				
WAIS-III PIQ	80.98 (13.15)	-12.62 (13.55)*	-6.72	51	0.60*
WAIS PIQ	93.60 (16.35)				
WAIS-III FSIQ	83.65 (17.46)	10.58 (15.15)*	6.13	76	0.54*
RAVEN	73.06 (13.11)				
WAIS-III FSIQ	79.61 (19.23)	13.49 (12.03)*	-6.44	32	0.79*
sGIT	93.09 (17.05)				

* $p < 0.001$.

(2%) from intellectual disability to normal. For the WAIS-III/WAIS comparison, 30 of the 62 cases (48%) changed category: 20 (32%) from borderline to normal, six (10%) from intellectual disability to borderline and three (5%) from intellectual disability to normal. The WAIS-III/ RAVEN and WAIS-III/sGIT categorical difference comparisons reached trend significance (Table 4). For the WAIS-III/RAVEN comparison, 47 of the 77 cases (61%) changed category: 17 (22%) from borderline to normal, 19 (25%) from intellectual disability to borderline and 11 (14%) from intellectual disability to normal. For the WAIS-III/sGIT comparison, 18 of the 33 cases (55%) changed category: 10 (30%) from borderline to normal, one (3%) from intellectual disability to borderline and seven (21%)

from intellectual disability to normal. Changes in category were not associated with time of administration, for example changes from normal to intellectual disability were not associated with longer duration between tests (results available upon request).

Table 4. Cross tabulation categories of IQ

	WAIS-III[2]			WAIS			RAVEN			sGIT		
WAIS-III[1]	≥85	71–84	≤70	≥85	71–84	≤70	≥85	71–84	≤70	≥85	71–84	≤70
≥85	14	0	0	22	3	0	10	14	11	7	0	0
71-84	5	18	0	17	6	3	3	11	14	10	7	0
≤70	1	9	8	3	3	4	0	5	9	7	1	1
Fisher exact		.00			.00			.07			.07	

Note. Fisher Exact *p*-values are reported.

Regression analyses

The WAIS-III[1] IQ scores significantly predicted the WAIS-III[2], WAIS, and RAVEN IQ scores (*p*-values < .001). Adding year of administration as a covariate to the model did not significantly change the direction of effect nor the *p*-values. Stratified analyses revealed that among participants with a history of special needs education or a diagnosis of intellectual disability, WAIS-III IQ scores did not significantly predict WAIS or RAVEN IQ scores (Table 5).

DISCUSSION

The stability and/or exchangeability of IQ scores was investigated in a forensic psychiatric sample with and Table 3 correlations between tests and IQ scores on one test significantly predicted scores on the other IQ tests, suggesting good stability between scores. However, when looking separately at individuals with a diagnosis of intellectual disability or history of special needs education, the stability between scores disappeared. In these individuals, a significant association between two IQ scores was only established when comparing the two WAIS-III scores. Furthermore, all comparisons between tests revealed significant differences between scores, with mean absolute differences larger than 10 points when comparing WAIS-III full-scale IQ and

Table 5. Multilevel regression analyses

		B	*p*	95% CI	*n*
WAIS-III[1] FSIQ	WAIS-III[2] FSIQ	0.92	.00	0.78–1.07	55
	Normal education	0.86	.00	0.64–1.08	35
	Special needs education	0.87	.00	0.65–1.09	20
	No diagnosis of ID	0.85	.00	0.65–1.05	40
	Diagnosis of ID	0.68	.00	0.35–1.01	15
	WAIS FSIQ	0.79	.00	0.60–0.98	61
	Normal education	0.59	.00	0.39–0.79	51
	Special needs education	0.03	.95	-0.80–0.86	7
	No diagnosis of ID	0.58	.00	0.38–0.79	48
	Diagnosis of ID	0.21	.51	-0.41–0.83	13
	RAVEN	0.59	.00	0.33–0.84	77
	Normal education	0.57	.00	0.25–0.90	53
	Special needs education	0.36	.15	-0.13–0.84	22
	No diagnosis of ID	0.53	.00	0.23–0.82	56
	Diagnosis of ID	0.01	.97	-0.56–0.54	21
	sGIT	0.07	.28	-0.06–0.19	33
	Normal education	0.41	.00	0.33–0.49	22
	Special needs education	0.38	.02	0.5–0.71	11
	No diagnosis of ID	0.03	.58	-0.07–0.14	24
	Diagnosis of ID	-0.08	.80	-0.72–0.55	9
WAIS-III[1] VIQ	WAIS-III[2] VIQ	0.97	.00	0.87–1.07	43
	Normal education	0.92	.00	0.80–1.05	28
	Special needs education	1.06	.00	0.83–1.30	15
	No diagnosis of ID	0.93	.00	0.80–1.05	31
	Diagnosis of ID	0.84	.00	0.45–1.22	12
WAIS-III[1] PIQ	WAIS-III[2] PIQ	0.89	.00	0.74–1.05	43
	Normal education	0.98	.00	0.82–1.15	28
	Special needs education	0.57	.00	0.30–0.83	15
	No diagnosis of ID	0.96	.00	0.78–1.14	31
	Diagnosis of ID	0.60	.00	0.32–0.87	12

performance scale with the WAIS full-scale IQ and performance scale and when comparing the WAIS-III with the RAVEN and sGIT. Frequencies of difference scores between tests also show substantial dissimilarities between tests, with a percentage of cases having more than a 10-point difference between tests ranging from 33% to 66%. These percentages are comparable to a study

investigating IQ stability in a forensic psychiatric (Van Toorn & Bon, 2011). In contrast, these percentages are much higher than those reported in the meta-analysis by Whitaker (2008b) investigating stability coefficients in individuals with low IQ (14%). The range of differences between scores is surprising, but the fact that there are differences is not. As mentioned in the introduction, IQ scores are not expected to have perfect temporal and instrumental stability, and there are several possible explanations for the differences in IQ scores. Certain factors such as dietary changes (Bellisle, 2004; Koyama et al., 2012; Smithers et al., 2012) and changes in quality of education or intellectual stimulation can result in a true change of IQ.

Chance error

All psychometric instruments are influenced by error and such is the case with the assessment of intelligence. Several sources of error are possible which can be classified in two broad categories: chance error and systematic error (Whitaker, 2010). Examples of chance error are: fluctuations in test performance or examiner's behavior, cooperation of the test subject and other personal and environmental factors. In a forensic psychiatric population, cooperation of the test subject can likely have a larger effect than expected in a non-forensic psychiatric population. It is possible that an individual intentionally performs worse to avoid prison (malingering) or is not motivated enough during the assessment as a result of his/hers psychiatric profile. Furthermore, the added stress of being arrested and sent to prison can result in lower scores. For example, Biles (1968) found significantly lower IQ scores upon arrival in prison compared to IQ scores obtained at a later time point during imprisonment. However, other research, albeit with a different approach, has found no effect of long-term imprisonment on IQ (Banister, Smith, Heskin, & Bolton, 1973; Bolton, Smith, Heskin, & Banister, 1976; Dettbarn, 2012; Goethals, 1981). Administration of an IQ test within a prison setting can also influence test scores. For example, obstacles such as a lack of privacy and adequate space, scheduling conflicts, and noise pollution could have an impact on test scores.

Systematic error: Flynn effect

An example of systematic error is the Flynn effect. The Flynn effect refers to the observation (Flynn, 1984) that every restandardization sample for a major intelligence test resulted in an IQ

score increase of approximately 0.33 points per year. The Flynn effect seems most prominent in people at the lower end of the distribution and in RAVEN scores (Colom, Lluis-Font, & Andres-Pueyo, 2005; Teasdale & Owen, 1989; Williams, 2013). For example, Teasdale and Owen (2005) found that the Flynn effect primarily reduced the number of low-end scores, resulting in an increased number of moderately high scores, with no increase in very high scores. In contrast, some studies have found a reverse Flynn effect with declining scores for those pursuing higher academic education and those not doing so (Dutton & Lynn, 2013; Teasdale & Owen, 2008).

In the present study, correction for the Flynn effect for the WAIS-III was not possible due to lack of information concerning which norms were used to calculate the WAIS-III scores. RAVEN scores from the present study were transformed using the latest Belgian norms available (Magez et al., 2006), thereby reducing potential increases as a result of the Flynn effect. In addition, due to the random sampling of scores, time of administration is balanced between subjects, again minimizing the potential impact of the Flynn effect in these analyses. Furthermore, although the increases found in this study are much larger than would be expected as a result of the Flynn effect alone, the Flynn effect could explain some of the differences in scores. Flynn stated that differences in scores over time do not reflect changes in that person's true IQ score, rather the differences are a result norms change (Flynn, 2006).

In contrast, researchers have also stated that because the Flynn effect concerns a rise in average IQ when comparing generations it does not apply to within-subject test–retest reliability (Rodgers, 1998; van Winkel et al., 2006).

The current study investigated differences between scores when the same individual is given the same test (WAIS-III$^{(1)}$ versus WAIS-III$^{(2)}$) and if the same individual is given different tests (e.g., WAIS-III versus RAVEN). Therefore it should be noted that differences in scores can have different causes in the former than in the latter. When comparing scores within the same instrument, changes are mainly due to chance error Whitaker (2010), whereas when comparing two instruments, both chance error and systematic error could result in changes in scores. Other examples of systematic error, that is floor effect and differences between IQ scales, are discussed in detail in Whitaker (2010).

Change due to mental disorder

In psychotic disorder, there is much debate about a potential progressive decline in cognitive functioning (Heaton et al., 2001; Zampera, 1999). In a 10-year follow-up study, pre-morbid IQ and post-morbid WAIS scores were compared in first episode patients with psychotic disorder. The results showed that patients with high premorbid IQs (≥108) had a 10-point decline in cognitive functioning; however, a restoration to pre-morbid level was found at follow-up (an average of 10 years later). In the low pre-morbid IQ group, a stable course of IQ was found (van Winkel et al., 2006). In the current sample, 18% of the participants had a diagnosis of psychotic disorder, which could explain some of the differences between scores. Unfortunately, it was not possible to investigate differences in stability of IQ scores between diagnoses in the present study due to a lack of power. Potential fluctuations of IQ due to a specific mental disorder should be taken into account when interpreting test results of a forensic psychiatric patient.

Thus, variation in scores may or may not represent the individual's true level of intellectual functioning. The term standard error of measurement is used to capture this variability and to provide a statistical confidence interval (CI) within which it is expected the individual's true score falls. Therefore, it is considered good practice to report CIs together with the full-scale IQ score. Most IQ tests report CIs of approximately 10 points (i.e., five points below or above the true IQ) (Whitaker, 2008a). For example, an individual's score of 70 on the WAIS-III corresponds with a 95% CI of 67–75 (Wechsler, 2005). The present study found absolute differences of more than 10 points in 18 of the 55 cases when comparing two WAIS-III scores and in 51 of the 77 cases when comparing the WAIS-III with the RAVEN, without even taking into account the level of education or diagnosis of intellectual disability. Therefore, depending on the test used, in 33% of the cases or even in 66% of the cases, the person's second score did not fall within his or her reported CI for the first score. This raises some implications for the interpretation of CIs in psychological reports.

Stability of IQ

In the previous sections, several explanations are given for the differences found in this study. However, these explanations do not alter the fact that disparity between test scores needs to be kept as minimal as possible, especially given the large consequences of inconsistent assessments of cognitive abilities for a forensic psychiatric patient. Intelligence does seem to be fairly stable

across the lifespan (Deary, Whiteman, Starr, Whalley, & Fox, 2004; Gow et al., 2011). Deary et al. (2004) investigated old intelligence scores from a sample of 90,000 Scottish children at ages 10 and 11 and reassessed them at the age of 80. They found a positive correlation of .66 between the two scores. However, this level of stability cannot simply be presumed in individuals with intellectual disability. Silverman et al. (2010) compared Stanford-Binet scores with WAIS scores in 74 individuals with intellectual disability. They found that when using the Stanford-Binet scores, 95% of their sample met the criteria for benefits through the Social Security Administration. In contrast, when using the WAIS scores, only 61% of the participants met the same criteria, resulting in a large number of individuals failing to comply with the criteria although their diagnosis of intellectual disability was already established and documented. Similar results are found in the present study. Depending on the type of test used, some individuals are classified as having an intellectual disability or are considered to have a normal IQ.

The American Association on Intellectual and Developmental Disabilities (AAIDD) postulates that a psychometric instrument performs best when used with individuals who score within 23 standard deviations of the mean (Schalock et al., 2010, p 39), whereas Whitaker (2013) states that IQ test perform reasonably well within one SD. Individuals with the diagnosis of intellectual disability fall in the extreme left tail of the IQ distribution (e.g., 2–3 standard deviations below the mean). It is therefore not surprising that the associations between IQ scores found in the present study disappeared when the present authors stratified on the basis of intellectual disability diagnosis and educational level. Nevertheless, the consequences of unstable IQ measurements can be great. As a rule, false positive diagnoses are expected to be rare because intellectual disability should not be diagnosed solely on the basis of IQ score. Three criteria need to be met before diagnosing intellectual disability: (i) significant limitations in intellectual functioning, (ii) significant limitations in adaptive behavior and (iii) age of onset before the age of 18 (American Psychiatric Association, 2000). In addition, an IQ score should not be viewed in isolation but should always be interpreted using environmental context, educational history and functioning of adaptive behavior. However, no safeguard exists for a false negative. The problems faced by individuals who have intellectual disability but do not receive the diagnosis of intellectual disability can be significant and the risk of a missed diagnosis is even higher in the people who fall within the borderline category. The current study showed significant differences in IQ

categories when using different IQ tests and even when using the same test at different times. Implementing an unified model of cognitive abilities in diagnostics could aid in avoiding false negatives or incorrect diagnoses and would help finding a better alignment between treatment and disabilities. A widely cited unified model on cognitive abilities is the Cattell–Horn–Carroll (CHC) model. This model is an empirically based model that approaches cognition as a multifactorial concept and is regarded as one of the most well-validated hierarchical taxonomies to classify and describe human cognitive abilities (McGrew, 2009). The CHC model could help to better disentangle learning disabilities, language disorders and intellectual disability and to fine tune treatments by focusing on the individuals' strengths and weaknesses based on his/hers CHC profile (e.g., Niileksela & Reynolds, 2014; Proctor, 2012). Examples on how application of the CHC model can benefit diagnosis in and treatment of individuals with limited cognitive abilities are described in Uzieblo et al. (2012) and Fiorello & Primerano (2005). Furthermore, new or adjustments to intelligence tests are increasingly using the CHC model as framework, as can be seen for instance in the newest version of the WAIS, the WAIS-IV.

Methodological considerations

Although information on time of administration was available, increases or decreases in IQ scores over time were not investigated in the current study. If a potential temporal association was investigated between scores, the samples sizes would have become too small to draw tangible conclusions out of the results. Larger longitudinal studies (in forensic psychiatric patients) are needed to further investigate which factors are responsible for the temporal changes found between scores. Also, it would be interesting to investigate exploratory factors other than the Flynn effect. IQ measurements were entered randomly in the data set with regard to time of administration and analyzed using that order. For example, when two WAISIII scores were available, it was possible that the WAISIII[(2)] score was an older score than WAIS-III[(1)]. Due to the random sampling of scores, no conclusion could be made in the current sample regarding increases in IQ scores over time as a result of the Flynn effect.

The practice effect refers to an increase in IQ score that results from an individual being retested on the same instrument (Kaufman, 1994). Therefore, established clinical practice is to avoid administering the same intelligence test within the same year to the same individual

because it will often lead to an overestimation of a person's true intelligence (Kaufman & Lichtenberger, 2006; Schalock et al., 2010). However, in court proceedings it is possible that an individual is being retested within a short time period by several experts. In addition, research has shown that people with lower IQ's have less *benefit* from practice effects (Rapport, Brines, Axelrod, & Theisen, 1997). In the current study, a number of IQ tests were readministered within a year, but the order of administration of each test is random, thereby averaging out possible practice effects.

CONCLUSION

The current study showed that although IQ scores are correlated within persons, stability and/or interchangeability of scores is lacking, especially in individuals with a great need for a stable assessment of intelligence (i.e., individuals with intellectual disability). Differences of 10 points and more were found between IQ assessments, with the largest differences found comparing the WAIS-III with the sGIT. Therefore, although current good practices entail reporting the confidence interval together with the IQ score, further caution in interpreting IQ scores is recommended. Additionally, all neuropsychological reports should contain information regarding the norms used and report raw scores. Uniformity in the use and reporting of intelligence measurements in forensic psychiatric patients is clearly necessary. The CHC model may serve as an important framework.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Banister, P. A., Smith, F. V., Heskin, K. J., & Bolton, N. (1973). Psychological correlates of long-term imprisonment: I Cognitive variables. *British Journal of Criminology, 13*, 312-323.

Bellisle, F. (2004). Effects of diet on behaviour and cognition in children. *British Journal of Nutrition, 92*(Suppl 2), S227-232.

Biles, D. (1968). Test performance and imprisonment. *Australian & New Zealand Journal of Criminology, 1*(1), 46-58.

Bolton, N., Smith, F. V., Heskin, K. J., & Banister, P. A. (1976). Psychological correlates of long-term imprisonment: IV A longitudinal analysis. *British Journal of Criminology, 16*, 38-47.

Colom, R., Lluis-Font, J. M., & Andres-Pueyo, A. (2005). The generational intelligence gains are caused by decreasing variance in the lower half of the distribution: Supporting evidence for the nutrition hypothesis. *Intelligence, 33*(1), 83-91. doi: doi:10.1016/j.intell.2004.07.010

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish mental surveys of 1932 and 1947. *Journal of Personality and Social Psychology, 86*(1), 130-147. doi: 10.1037/0022-3514.86.1.130

Dettbarn, E. (2012). Effects of long-term incarceration: A statistical comparison of two expert assessments of two experts at the beginning and the end of incarceration. *International Journal of Law and Psychiatry, 35*(3), 236-239. doi: 10.1016/j.ijlp.2012.02.014

Nuovo, A. G. Di, Nuovo, S. Di, & Buono, S. (2012). Intelligent Quotient estimation of mental retarded people from different psychometric instruments using artificial neural networks. *Artificial Intelligence in Medicine, 54*(2), 135-145. doi: 10.1016/j.artmed.2011.11.002

Dutton, E., & Lynn, R. (2013). A negative Flynn effect in Finland, 1997–2009. *Intelligence, 41*, 817-820.

Emmerik, J. L. van (2001). *De terbeschikkingstelling in maat en getal.* Den Haag: Ministerie van Veiligheid en Justitie, DJI.

Evans, I. M. (1991). Testing and diagnosis: A review and evaluation. In L. H. Meyer, C. A. Peck, & L. Brown (Eds.), *Critical issues in the lives of people with severe disabilities*. Baltimore, MD: Brookes Publishing Co.

Fabian, J. M., Thompson, W. W., & Lazarus, J. B. (2011). Life, death, and IQ: It's much more than just a score: Understanding and utilizing forensic psychological and neuropsychological evaluations in Atkins intellectual disability/mental retardation cases. *Cleveland State Law Review, 59*, 402-430.

Fiorello, C. A., & Primerano, D. (2005). Research into practice: Cattell-Horn-Carroll cognitive assessment in practice: Eligibility and program development issues. *Psychology in the Schools, 42*(5), 525-536. doi: 10.1002/pits.20089

Floyd, R. G., Clark, M. H., & Shadish, W. R. (2008). The exchangeability of IQs: Implications for professional psychology. *Professional Psychology: Research and Practice, 39*, 414-423.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 101*, 171-191.

Flynn, J. R. (2006). Tethering the elephant: Capital cases: IQ, and the Flynn effect. *Psychology, Public Policy, and Law, 12*(2), 170-189.

Goethals, J. (1981). Imprisonment and cognitive impairment. *Journal of Clinical Psychology, 37*(2), 418-422.

Gow, A. J., Johnson, W., Pattie, A., Brett, C. E., Roberts, B., Starr, J. M., & Deary, I. J. (2011). Stability and change in intelligence from age 11 to ages 70, 79, and 87: The Lothian Birth Cohorts of 1921 and 1936. *Psychology and Aging, 26*(1), 232-240. doi: 10.1037/a0021072

Heaton, R. K., Gladsjo, J. A., Palmer, B. W., Kuck, J., Marcotte, T. D., & Jeste, D. V. (2001). Stability and course of neuropsychological deficits in schizophrenia. *Archives of General Psychiatry, 58*(1), 24-32.

Jensen, A. R. (1980). *Bias in mental testing.* London: Methuen.

Kaufman, A. S. (1994). Practice effects. In R. J. Sternberg (Ed.), *Encyclopedia of intelligence* (Vol. II, pp. 828-833). New York, NY: Macmillan.

Kaufman, A. S., & Kaufman, N. L. (1993). *Kaufman Adolescent and Adult Intelligence Test (KAIT) manual*. Circle Pines, MN: American Guidance Service, Inc.

Kaufman, A. S., & Lichtenberger, E. O. (2006). *Assessing adolescent and adult intelligence*. Needham Heights, MA Allyn & Bacon.

Kooreman, A., & Luteijn, F. (1987). *Groninger Intelligentie Test: schriftelijke verkorte vorm*. Lisse: Swets & Zeitlinger.

Koyama, K. I., Asakawa, A., Nakahara, T., Amitani, H., Amitani, M., Saito, M., . . . Inui, A. (2012). Intelligence quotient and cognitive functions in severe restricting-type anorexia nervosa before and after weight gain. *Nutrition, 28*(11-12), 1132-1136. doi: 10.1016/j.nut.2012.03.003

Luteijn, F., & Barelds, D. P. (2004). *Git2. Groninger Intelligentie Test 2.* Amsterdam: Harcourt Assessment.

Magez, W., Moenaert, H., & Degezelle, A. (2006). RAVEN Standaard Progressive Matrices (SPM) Normen voor volwassen in Vlaanderen in relatie met de WAIS-III Nederlandstalig bewerking. Leuven: CAPvzw.

McBrien, J. (2003). The intellectually disabled offender: Methodological problems in identification. *Journal of Applied Research in Intellectual Disabilities, 16*(2), 95–105. doi: 10.1046/j.1468-3148.2003.00153.x

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*(1), 1-10. doi: 10.1016/j.intell.2008.08.004

Moenaert, H. (2006). *WAIS-III NL en Raven SPM bij Vlaamse arbeidsgehandicapten*. Leuven Federatie Consultatiebureaus België.

Niileksela, C. R., & Reynolds, M. R. (2014). Global, broad, or specific cognitive differences? Using a MIMIC model to examine differences in CHC abilities in children with learning disabilities. *Journal of Learning Disabilaties, 47*(3), 224-236. doi: 10.1177/0022219412453172

Proctor, B. (2012). Relationships between Cattell-Horn-Carroll (CHC) cognitive abilities and math achievement within a sample of college students with learning disabilities. *Journal of Learning Disabilaties, 45*(3), 278-287. doi: 10.1177/0022219410392049

Rapport, L. J., Brines, D. B., Axelrod, B. N., & Theisen, M. E. (1997). Full scale IQ as mediator of practice effects: The rich get richer. *The Clinical Neuropsychologist, 11*(4), 375-380.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.

Raven, J. C., Court, J. H., & Raven, J. (1983). *Manual for Raven's progressive Matrices and Vocabulary Scales*. London: Lewis.

Rodgers, J. L. (1998). A critique of the Flynn effect: Massive IQ gains, methodological artifacts, or both? *Intelligence, 26*(4), 337-356. doi: 10.1016/S0160-2896(99)00004-5

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales* (5th ed.). Itasca, IL: Riverside Publishing.

Schalock, R. L., Borthwick-Duffy, S. A., Bradley, V. J., Buntinx, W. H. E., Coulter, D. L., Craig, E. M., . . . Yeager, M. H. (2010). *Intellectual Disability: Definition, Classification, and Systems of Supports* (11th ed.). Washington, D.C.: American Association on Intellectual and Developmental Disabilities

Schroth, M. L. (1983). A study of aging, intelligence and problem solving. *Psychological Reports, 53*, 1271-1279. doi: 10.2466/pr0.1983.53.3f.1271

Silverman, W., Miezejeski, C., Ryan, R., Zigman, W., Krinsky-McHale, S., & Urv, T. (2010). Stanford-Binet & WAIS IQ differences and their implications for adults with intellectual disability (aka mental retardation). *Intelligence, 38*(2), 242-248. doi: 10.1016/j.intell.2009.12.005

Smithers, L. G., Golley, R. K., Mittinty, M. N., Brazionis, L., Northstone, K., Emmett, P., & Lynch, J. W. (2012). Dietary patterns at 6, 15 and 24 months of age are associated with IQ at 8 years of age. *European Journal of Epidemiology, 27*(7), 525-535. doi: 10.1007/s10654-012-9715-5

StataCorp. (2011). *Stata Statistical Software: Release 12. College Station*. TX: StataCorp LP.

Teasdale, T. W., & Owen, D. R. (1989). Continuing secular increases in intellgence and a stable prevalence of high intelligence levels. *Intelligence, 13*, 255-262.

Teasdale, T. W., & Owen, D. R. (2005). A long term rise and recent decline in intelligence test performance: The Flynn Effect in reverse. *Personality and Individual Differences, 39*, 837-843. doi: doi:10.1016/j.paid.2005.01.029

Teasdale, T. W., & Owen, D. R. (2008). Secular declines in cognitive test scores: A reversal of the Flynn effect. *Intelligence, 36*(2), 121-126. doi: 10.1016/j.intell.2007.01.007

Tellegen, P. (2002). De kwaliteit van de normen van de WAIS-III [The quality of the WAIS-III norms]. *De Psycholoog, 37*, 463-465.

Toorn, B. Van, & Bon, C. (2011). De onbetrouwbaarheid van de IQ-meting [The unreliability of IQ-tests]. *De Psycholoog, 4*, 44-49.

Tulkin, S. R., & Newbrough, J. R. (1968). Social class, race, and sex differences on the Raven. *Journalf of Consulting and Clinical Psychology, 32*, 400-406.

Uzieblo, K., Winter, J., Vanderfaeillie, J., Rossi, G., & Magez, W. (2012). Intelligent diagnosing of intellectual disabilities in offenders: Food for thought. *Behavioral Sciences & the Law, 30*(1), 28-48. doi: 10.1002/bsl.1990

Verlinden, S., Maes, B., & Goethals, J. (2009). *Personen met een verstandelijke handicap onderhevig aan een interneringsmaatregel*. Leuven: Steunpunt Welzijn, Volksgezondheid en Gezin.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale WAIS*. New York, NY: Psychological Corporation.

Wechsler, D. (1970). *WAIS Nederlandstalige bewerking.Technische Handleiding* Amsterdam: Swets & Zeitlinger.

Wechsler, D. (1997). *The Wechsler Adult Intelligence Scale – Third Edition*. San Antonio, TX: Psychological Corporation

Wechsler, D. (2000). *WAIS-III Nederlandstalige bewerking.Technische Handleiding*. Lisse: Swets & Zeitlinger.

Wechsler, D. (2005). *WAIS-III Nederlandstalige bewerking. Technische handleiding*. Lisse: Swets & Zeitlinger.

Whitaker, S. (2008a). Intellectual disability: A concept in need of revision? *The British Journal of Developmental Disabilities, 54*(106), 3-9.

Whitaker, S. (2008b). The stability of IQ in people with low intellectual ability: An analysis of the literature. *Intellectual and Developmental Disabilities, 46*(2), 120-128. doi: 10.1352/0047-6765(2008)46[120:TSOIIP]2.0.CO;2

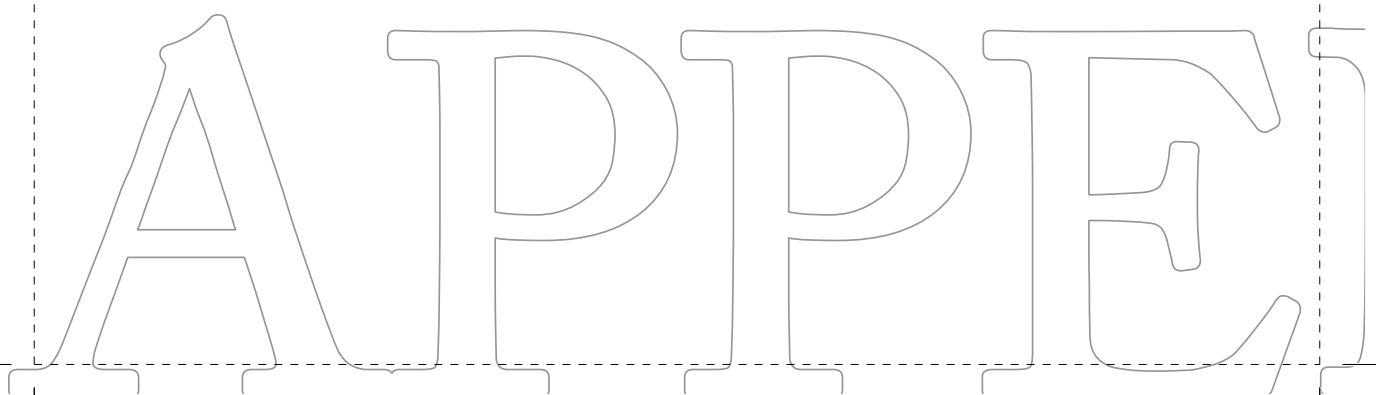
Whitaker, S. (2010). Error in the estimation of intellectual ability in th elow range using WISC-IV and WAIS-III. *Personality and Individual Differences, 48*(5), 517-521. doi: doi:10.1016/j.paid.2009.11.017

Whitaker, S. (2013). *Intellectual Disability: An Inability to Cope with an Intellectually Demanding World*. London: Palgrave MacMillan.

Williams, R. L. (2013). Overview of the Flynn effect. *Intelligence, 41*(6), 753-764. doi: 10.1016/j.intell.2013.04.010

Winkel, R. van, Myin-Germeys, I., Delespaul, P., Peuskens, J., Hert, M. De, & Os, J. van (2006). Premorbid IQ as a predictor for the course of IQ in first onset patients with schizophrenia: A 10-year follow-up study. *Schizophrenia Research, 88*(1-3), 47-54. doi: 10.1016/j.schres.2006.06.033

Zampera, E. (1999). Intellectual performance of chronic schizophrenic patients. *Colegiuml Antropologicum*, *23*(2), 597-602.



De IQ-score in twijfel getrokken: Evolutie naar een multidimensioneel cognitief vaardigheidsprofiel

Een illustratie uit de forensische psychiatrie

Uzieblo, K., Habets, P., & Jeandarme, I.

Neuropraxis, 19(4), 87-94

APPENDIX

SAMENVATTING

Een foutieve of ontoereikende inschatting van het intellectueel functioneren kan verstrekkende gevolgen hebben binnen de forensische context. Niet zelden worden beschermings- en strafmaatregelen afgestemd op de intellectuele vaardigheden van de pleger. Een kwalitatieve bepaling van dit functioneren is dan ook van groot belang. Dit blijkt echter geen vanzelfsprekend gegeven te zijn. Een recente studie naar de stabiliteit van herhaalde IQ-metingen bij Vlaamse forensische patiënten (*n* = 176) (Habets, Jeandarme, Uzieblo, Oei, & Bogaerts, 2014) geeft aan dat verschillende intelligentietesten en herhaalde metingen van eenzelfde intelligentietest niet steeds gelijkaardige resultaten opleveren. Scores verkregen op basis van intelligentietesten blijken dus niet zonder meer inwisselbaar te zijn. De huidige studie beoogt een kritische blik te werpen op de waarde van deze intelligentiemetingen en op de recente ontwikkelingen binnen dit veld. Meer specifiek worden de mogelijke oorzaken van de instabiliteit van IQ-scores besproken. Aansluitend worden de belangrijkste ontwikkelingen aangaande de diagnostiek van het intellectueel functioneren uiteengezet. Een gewijzigd denkkader treedt de laatste jaren immers meer en meer op de voorgrond, waarbinnen de klemtoon stelselmatig van een IQ-score naar het in kaart brengen van een breed cognitief vaardigheidsprofiel en het adaptief functioneren verschuift.

INLEIDING

Ik werk als psycholoog in de gevangenis met geïnterneerden met een verstandelijke handicap. In expertiseverslagen wordt er vaak gebruik gemaakt van de Progressieve Matrices van Raven. Later in het traject wordt er soms nog een WAIS-III afgenomen en de resultaten op deze testen kunnen zeer uiteenlopend zijn. Zo hebben we bij een bepaalde patiënt bijvoorbeeld een IQ-score van 54 met de Progressieve Matrices verkregen, terwijl de WAIS-III dan weer een IQ-score van 88 aangeeft. Dit is nu één voorbeeld, maar dergelijke over- of onderschattingen komen we vaak tegen. Ik las dat er een vrij sterke correlatie bestaat tussen de Progressieve Matrices en de WAIS-testen (.40 tot .75). Momentane invloeden spelen wellicht ook een rol (bijvoorbeeld weerstand, onwil,…). Kunt u mij hier wat informatie rond bezorgen?

Bovenstaande vraag trekt de eenduidigheid en stabiliteit van intelligentiemetingen in twijfel, twee pijnpunten binnen onze testdiagnostiek die tot op heden maar beperkte aandacht gekregen hebben. Met dit artikel beogen we een kritische analyse te geven van de huidige intelligentiemetingen, gekaderd binnen de forensische praktijk. Daarnaast bespreken we de recente evolutie naar een multidimensionele benadering van het intellectueel (dis)functioneren.

INTELLIGENTIE EN ANTISOCIAAL GEDRAG

Naar schatting zou 1% van de wereldbevolking een intellectuele disfunctie[1] hebben (Maulik, Mascarenhas, Mathers, Dua, & Saxena, 2011). Binnen het strafrechtsysteem doen we echter een verontrustende ontdekking: het voorkomen van een intellectuele disfunctie bij daders wordt geschat op 2-10%, afhankelijk van de instelling (bijvoorbeeld gevangenis, forensische psychiatrie, enz.) en het land (Lindsay, 2011). Ook in België ligt het voorkomen van een intellectuele disfunctie bij daders relatief hoog. Zo wordt bij 4% van de geïnterneerden die zijn opgenomen in het psychiatrisch zorgcircuit van de Federale Overheidsdienst Volksgezondheid (o.a., psychiatrische verzorgingstehuizen, de ambulante hulpverlening, instellingen tot de bescherming van de maatschappij, beschut wonen, medium security units en afdelingen binnen de algemene psychiatrie) een intellectuele disfunctie (zoals gedefinieerd in eerdere DSM-edities; i.e., IQ < 70)

als hoofddiagnose gesteld in vergelijking tot 17% van de geïnterneerden die zijn opgesloten in een strafinstelling (Cosyns, 2005). Daarnaast zou 15% van de geïnterneerden zwakbegaafd zijn, met IQ-scores tussen 70 en 85 (Maes, Goethals, & Verlinden, 2009). Men gaat ervan uit dat deze cijfers een onderschatting zijn: het crimineel gedrag dat wordt vertoond door personen met een intellectuele disfunctie zou onvoldoende systematisch aan de politie gerapporteerd worden (Jones, 2007).

Deze cijfers roepen vragen op omtrent de aard van de relatie tussen het intellectueel functioneren en antisociaal gedrag. Hoewel men deze relatie op zich niet meer in twijfel trekt, zijn er nog veel onduidelijkheden wat betreft de aard en de sterkte van de relatie. Er bestaat evidentie voor zowel een negatief verband als voor een curvilineair verband tussen IQ en antisociaal gedrag (Mears & Cochran, 2013). Doorgaans stelt men vast dat personen met IQ-scores onder de 50 maar zelden in contact komen met justitie, terwijl de groep met een lichte intellectuele disfunctie en de zwakbegaafde groep dan weer over gerepresenteerd zijn binnen de forensische populatie (Lindsay, Hastings, & Beail, 2013). Een hogere IQ-score wordt als protectieve factor beschouwd (Portnoy, Chen, & Raine, 2013). Bovendien kan men niet stellen dat IQ-scores toekomstig antisociaal gedrag voorspellen. Longitudinale studies tonen aan dat een hele reeks factoren, zoals leeftijd en het neurocognitief functioneren (Loeber et al., 2012), genetische factoren (DeYoung et al., 2006; Koenen, Caspi, Moffitt, Rijsdijk, & Taylor, 2006), en gedragsproblemen (Lahey, D'Onofrio, Van Hulle, & Rathouz, 2014) inspelen op de relatie tussen IQ-scores en de ontwikkeling van antisociaal gedrag. Dus in tegenstelling tot eerdere presumpties (Herrnstein & Murray, 1994), kunnen we niet stellen dat lage IQ-scores op zichzelf de ontwikkeling van antisociaal gedrag voorspellen. Eerder zal de vaak complexe interactie tussen het intellectueel functioneren en andere criminogene factoren de uitkomst bepalen.

DE IQ-SCORES IN TWIJFEL GETROKKEN

Een betrouwbare en valide bepaling van het intelligentieniveau is binnen de forensische context van groot belang. Bij een foutieve inschatting komen twee basisprincipes van justitie zwaar in het gedrang. Justitie dient immers te verzekeren dat de betrokkene de rechten, plichten en juridische procedures ten volle begrijpt en er ook gebruik van kan maken. Daarnaast dient een aangepaste

begeleiding en behandeling voor alle kwetsbare verdachten/slachtoffers/getuigen (o.a., drugsverslaafden, jongeren, en personen met psychische problemen) door politie en justitie verzekerd te worden. Een foutieve inschatting van het intelligentieniveau kan bijvoorbeeld valse verklaringen tijdens het verhoor in de hand werken (Kassin, 2008). Verder speelt het intelligentieniveau een grote rol bij het advies aangaande de toerekeningsvatbaarheid, de inschatting van de kans op recidive en de indicatiestelling (zie bijvoorbeeld Van Toorn & Bon, 2011). In sommige landen kan een klinisch lage IQ-score een dader zelfs behoeden voor een terdoodveroordeling (Appelbaum, 2014). Een onder- of overschatting van de intellectuele capaciteiten kan met andere woorden een fair verloop van het juridische proces ernstig bedreigen.

Intelligentiemetingen zijn echter niet feilloos of zonder meer inwisselbaar; zowel de instrumentele en temporele stabiliteit van IQ-scores worden in twijfel getrokken (Evans, 1991). De diagnose van een intellectuele disfunctie is in belangrijke mate afhankelijk van welk instrument gebruikt werd om intelligentie te bepalen (Casey & Keilitz, 1990; Van Toorn & Bon, 2011) met alle gevolgen van dien. Ondanks het feit dat scores tussen IQ-testen op groepsniveau gecorreleerd zijn, blijken scores binnen personen vaak niet identiek te zijn (Di Nuovo, Di Nuovo, & Buono, 2012; Floyd, Clark, & Shadish, 2008) en dit verschil is nog groter bij mensen met een intellectuele disfunctie (Silverman et al., 2010; Whitaker, 2010). In een recente studie werd de instrumentele en temporele stabiliteit van herhaalde IQ-metingen in een forensische populatie (met en zonder intellectuele disfunctie) nagegaan (Habets, Jeandarme, Uzieblo, Oei, & Bogaerts, 2014). Deze studie maakte deel uit van een grote Vlaamse observationele studie naar recidivecijfers bij forensische patiënten. Tijdens de dossierstudie werden van 167 patiënten (97% mannen) dubbele IQ-scores gevonden van de Nederlandse versies van volgende intelligentietesten: de Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955, 1970), de WAIS-III[2] (Wechsler, 1997; Wechsler 2005), Raven's Progressieve Matrices (RAVEN; Raven, Raven, & Court, 1998) en de verkorte Groninger Intelligentie Test (vGIT; (Kooreman & Luteijn, 1987)). De langste gemiddelde tijd tussen twee afnames was, zoals verwacht, tussen de WAIS-III en de WAIS (8.6 jaar). Voor de andere vergelijkingen waren de verschillen 1.4, 3.7 en 1.3 jaar voor de vergelijking met de andere WAIS-III score, de RAVEN en de vGIT respectievelijk. Zoals verwacht, werden hoge correlaties tussen de intelligentietesten gevonden, met de hoogste correlatie tussen de verbale subtesten

van twee WAIS-III-afnamen (r = .96) en de laagste (doch significante) correlatie tussen de totale WAIS-III score en de score verkregen op de RAVEN (r = .54). Ook voorspelde de IQ-score van een test die van een andere test, hetgeen een goede stabiliteit tussen scores suggereert. Maar wanneer er apart werd gekeken naar de patiënten met een intellectuele disfunctie was er van stabiliteit tussen scores maar weinig sprake meer. Zo voorspelden bijvoorbeeld de WAIS–III-scores de WAIS- en de RAVEN-scores voor de gehele populatie, maar niet voor personen met een intellectuele disfunctie. Ten aanzien van het gemiddelde verschil tussen instrumenten, werden er significante verschillen gevonden tussen testen, met zelfs gemiddelde verschillen groter dan 10 punten (bijvoorbeeld WAIS-III vs. RAVEN; WAIS-III vs. vGIT). Verschillen van een dergelijke orde van grootte kunnen dus betekenen dat bepaalde personen op het ene moment in de categorie *intellectueel disfunctie* vallen en op het andere moment geclassificeerd worden als normaal begaafd, wat ook in deze studie werd vastgesteld. Aan de hand van de IQ-scores werden de personen ingedeeld in drie categorieën: normaal begaafd ($IQ \geq 85$), zwakbegaafd ($70 < IQ < 85$) en zwakzinnigheid ($IQ \leq 70$). De indeling in deze categorieën was verschillend afhankelijk van welke test gebruikt werd. In 27% van de gevallen werd aan personen een andere categorie toegewezen wanneer er twee WAIS-III scores werden gebruikt. Dit percentage was nog hoger bij de vergelijkingen van de WAIS-III met de WAIS (48%) en met de Raven (61%).

Verschillen in IQ-scores kunnen meerdere oorzaken hebben. Bepaalde factoren zoals aanpassingen in eetgewoonten (Bellisle, 2004; Koyama et al., 2012; Smithers et al., 2012), beter onderwijs of meer intellectuele stimulatie kunnen resulteren in een verandering van het IQ. Maar ook meetfouten kunnen een rol spelen en kunnen ingedeeld worden in twee brede categorieën: toevalsfouten en systematische fouten (Whitaker, 2010). Voorbeelden van toevalsfouten zijn: fluctuaties in het gedrag van de onderzoeker, de mate van medewerking van de testpersoon en andere persoonlijke en omgevingsfactoren. Zo zal binnen de forensische context de mate van medewerking een niet te onderschatten rol spelen. Het is bijvoorbeeld mogelijk dat een persoon opzettelijk slechter presteert (*malingering*) om de gevangenis te vermijden of er is een gebrek aan motivatie tijdens de testafname vanwege het psychiatrisch profiel van de persoon in kwestie. Ook kan de stress die voortkomt uit een arrestatie en opsluiting ervoor zorgen dat een persoon slechter scoort. Belangrijke systematische fouten bij intelligentiemetingen zijn het vloereffect, leereffecten en het Flynn effect. Een vloereffect houdt in dat men aan de hand van de test

moeilijk een onderscheid kan maken tussen zwakke en zeer zwakke scores, omdat te veel personen een lage score op de test of subtest hebben behaald. Dit vloereffect zou meer een probleem vormen bij de Wechsler voor kinderen dan bij de Wechsler voor volwassenen (Whitaker, 2010), maar hier is nog maar heel weinig onderzoek naar gedaan. Een leereffect verwijst naar een scoreverbetering vanwege een herhaalde testafname en dient in aanmerking genomen te worden, aangezien plegers tijdens hun justitieel traject doorgaans aan meerdere intelligentiemetingen onderworpen worden. De (maar) enkele beschikbare studies suggereren dat intelligentie-instrumenten inderdaad onderhevig kunnen zijn aan leereffecten (Estevis, Basso, & Combs, 2012; Staff, Hogan, & Whalley, 2014; Wechsler, 2008, 2012). De vraag blijft echter over welke tijdsperiode deze leereffecten standhouden en in welke mate verschillende afnames cumulatieve effecten met zich meebrengen (zie bijvoorbeeld Staff et al., 2014). De derde belangrijke systematische fout is het Flynn effect, hetgeen verwijst naar de observatie van een verhoging van 0.33 punten per jaar na elke hernormering van een intelligentietest (Flynn, 1984; Trahan, Stuebing, Fletcher, & Hiscock, 2014). De hier beschreven fouten zouden de resultaten van Habets en collega's (2014) deels kunnen verklaren. Niettegenstaande werd echter in deze studie de tijd van afname van de IQ-testen gebalanceerd tussen proefpersonen door het feit dat de IQ-scores at random verzameld werden, waardoor de invloed van leereffecten en van het Flynn-effect werd geminimaliseerd.

De IQ-testen zijn daarenboven niet in alle doelgroepen even betrouwbaar te gebruiken. Hierbij denken we onder meer aan de psychiatrische populatie en personen met een intellectuele disfunctie. Zo kunnen IQ-scores in de loop van de tijd veranderen vanwege een psychiatrische aandoening. Bij de psychotische stoornis is er bijvoorbeeld veel discussie omtrent de progressieve achteruitgang van het cognitief functioneren (Heaton et al., 2001; Zampera, 1999). Onderzoek toonde aan dat patiënten met hoge premorbide scores (≥108) na hun eerste psychotische periode een daling van 10 punten lieten zien. Deze daling werd niet meer gemeten bij de follow-upmeting die gemiddeld tien jaar later werd gedaan: de IQ-scores waren namelijk weer vergelijkbaar met de premorbide scores (van Winkel et al., 2006). In de groep plegers met een intellectuele disfunctie vertoont doorgaans een aanzienlijk percentage een diagnose van een psychotische stoornis (Dias, Ware, Kinner, & Lennox, 2013; Habets et al., 2014), wat het belang hiervan nogmaals onderstreept. Ook de inname van psychiatrische medicatie kan een

significante wijziging in het cognitief (Levin et al., 2015) en intellectueel functioneren met zich meebrengen (Gimpel et al., 2005). Bij het interpreteren van IQ-scores van een forensisch patiënt moet men dus rekening houden met mogelijke schommelingen in IQ scores vanwege een specifieke psychische stoornis en medicatie-inname. Daarnaast werden de psychometrische kwaliteiten van de IQ-testen in twijfel getrokken wanneer deze gebruikt worden bij personen met een intellectuele disfunctie (Schalock et al., 2010, p 39). De factorstructuur van de Wechsler-schalen blijkt bijvoorbeeld immers niet stand te houden binnen deze groep (MacLean, McKenzie, Kidd, Murray, & Schwannauer, 2011; Reynolds, Ingram, Seeley, & Newby, 2013) en aan het lagere einde van het continuüm lijken de testen significant minder te kunnen differentiëren (MacLean et al., 2011).

Om de variabiliteit in IQ-scores enigszins op te vangen wordt het sterk aanbevolen om het betrouwbaarheidsinterval te vermelden bij de IQ-score. De meeste IQ-testen rapporteren betrouwbaarheidsintervallen van ongeveer 10 punten (5 boven en 5 onder de score) (Whitaker, 2008). Bijvoorbeeld een WAIS-III IQ-score van 70 heeft een 95% betrouwbaarheidsinterval van 67-75 (Wechsler 2005). Maar ook deze methodiek lost het probleem niet op. Uit de studie van Habets et al. (2014) bleek immers dat een aanzienlijk aantal scores zelfs buiten het betrouwbaarheidsinterval vallen. Afhankelijk van de test werd er in 33% (tot zelfs 66%) van de gevallen een tweede score gerapporteerd die niet binnen het betrouwbaarheidsinterval viel. Mede door deze problemen onderstreept men in de recente literatuur het belang van een multidimensionele benadering van het intellectueel (dis)functioneren.

EEN MULTIDIMENSIONEEL KADER VOOR HET INTELLECTUEEL FUNCTIONEREN

De multidimensionele benadering van intelligentie en de bijhorende diagnostiek wordt momenteel op twee domeinen onderstreept. Ten eerste wordt de focus nu meer dan ooit op het cognitieve vaardigheidsprofiel gelegd met een diepgaande differentiatie van de verschillende mogelijke cognitieve vaardigheden. Ten tweede treedt bij de beoordeling van een intellectuele disfunctie het adaptief functioneren op de voorgrond en dient men dus meerdere dimensies in kaart te brengen voordat een intellectuele disfunctie vastgesteld kan worden.

Verschillende IQ-testen hebben intussen de aanvankelijk unifactoriële (focus op het IQ-cijfer) of dichotome visie (focus op de verbale en non-verbale component) op intelligentie verlaten en zijn geëvolueerd naar een multifactoriële benadering. Deze evolutie is bijvoorbeeld merkbaar in de laatste WAIS-editie. Zo komt in deze editie de dichotome structuur volledig te vervallen en benadrukt men een geoptimaliseerde vierfactorenstructuur, bestaande uit een index voor verbaal begrip, perceptueel redeneren, het werkgeheugen en de verwerkingssnelheid. Uit de recente theorieën over intelligentie blijkt echter dat ook deze vierdeling nog ontoereikend is. Een van de meest geciteerde theorieën voor cognitieve vaardigheden, met name het Cattell-Horn-Carroll (CHC) model, wordt momenteel beschouwd als het meest omvangrijke en best empirisch gevalideerd model (McGrew, 2009). Het model is hiërarchisch geordend in drie niveaus of zogenaamde strata. De algemene *g-factor* bevindt zich op het derde niveau en is niet rechtstreeks te meten. Deze g-factor is verder opgebouwd uit verschillende brede cognitieve vaardigheden (BCV), die op het tweede niveau gesitueerd worden. Dit zijn vrij basale en stabiele cognitieve eigenschappen in een persoon, die gedrag sturen en beïnvloeden. Doorgaans worden de volgende BCV's onderscheiden: vloeiende intelligentie (Gf; o.a., het vermogen om relatief nieuwe taken op te lossen), kwantitatieve kennis (Gq; o.a., verworven wiskundige kennis en wiskundig redeneren), gekristalliseerde intelligentie (Gc; o.a., verworven kennis, afhankelijk van scholing en culturele ontwikkeling), lezen en schrijven (Grw), kortetermijngeheugen (Gsm; o.a., vaardigheid om informatie kort te onthouden en te verwerken), visuele informatieverwerking (Gv; o.a., vaardigheden om visuele patronen en prikkels waar te nemen, te analyseren, te synthetiseren en ermee te denken), auditieve informatieverwerking (Ga; o.a., vaardigheden om auditieve prikkels te begrijpen, te analyseren en te synthetiseren), langetermijngeheugen (Glr; o.a., vaardigheid waarmee informatie efficiënt wordt opgeslagen en de toegankelijkheid van die informatie) en verwerkingssnelheid (Gs; o.a., vaardigheid om eenvoudige cognitieve taken vloeiend en automatisch uit te voeren). De BCV's bestaan op hun beurt uit meer dan 70 nauwe cognitieve vaardigheden (NCV's), ofwel de bijkomende specialisaties die vereist zijn om informatie te verwerken binnen de specifieke domeinen van de BVC's (NCV, McGrew, 2009). De NCV's worden via de subtesten van de intelligentie-instrumenten gemeten. Ter illustratie, de WAIS-IV subtest *Gewichten* gaat de NCV inductie na, hetgeen deel uitmaakt van de BCV Gf. De premisse hierbij is dat hoe meer BCV's gemeten worden, hoe nauwkeuriger men de algemene intelligentie

g kan inschatten (zie voor een overzicht, Keith & Reynolds, 2010). Een gelijkaardige tendens vindt men terug in de huidige classificatiesystemen voor de definiëring van de intellectuele disfunctie. In de Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) neemt men afstand van de voorheen unidimensionele benadering met een sterke focus op het IQ-cijfer. Tekorten in de intellectuele vaardigheden worden nu aan de hand van tekorten op verscheidene cognitieve domeinen "als het redeneervermogen, de probleemoplossende vaardigheden, planning, abstract denken, oordeelsvermogen, academisch leren, en het leren uit ervaring" beschreven (American Psychiatric Association, 2013).

De bepaling van de intellectuele vaardigheden (en beperkingen) dient met andere woorden gebaseerd te zijn op een breed assessment in combinatie met het klinisch oordeel van de clinicus (American Psychiatric Association, 2013), en dus niet meer louter op basis van een intelligentietest. Greenspan en Woods (2014) opperen daarom dat een breed assessment enkel mogelijk gemaakt kan worden door gebruik te maken van neuropsychologische testen. Maar net als in de DSM ontbreekt het hier aan een theoretisch kader, waardoor de kans reëel blijft dat men belangrijke vaardigheden in de diagnostiek zal missen. Een theorie-gedreven diagnostiek is hier eerder aan te raden. In Vlaanderen is er echter momenteel geen enkele test die een theorie-gedreven (bijvoorbeeld CHC-gedreven) assessment van de cognitieve vaardigheden toelaat. Dit geldt ook voor de Wechsler-schalen die een duidelijk theoretisch kader missen. Indien men de Wechsler-schalen bijvoorbeeld binnen het CHC-kader evalueert, blijken de laatste edities belangrijke vaardigheden als auditieve informatieverwerking (Ga) en langetermijn geheugen (Glr) niet te meten (Flanagan, Ortiz, & Alfonso, 2013). Bovendien geeft een dergelijke theoretische evaluatie de toch wel frappante inhoudelijke verschillen tussen de courant gehanteerde intelligentie-instrumenten. Zo blijkt de totaalscore van de Raven een inschatting te geven van slechts een bepaalde component van het intellectueel functioneren, met name de Gf, in tegenstelling tot bijvoorbeeld de WAIS-III die Gf, Gc, Gv, Gsm en Gs in kaart brengt (zie ook, Alfonso et al., 2005). Dergelijke inhoudelijke verschillen kunnen dus tevens (deels) ten grondslag liggen aan verschillen tussen intelligentiescores die zijn verkregen op basis van verscheidene instrumentaria.

Een CHC-gedreven diagnostiek zou evenwel verscheidene voordelen met zich meebrengen. Ze biedt niet alleen een eenduidige nomenclatuur aan, ze geeft ook een kader voor onder meer

een diepgaande differentiaaldiagnostiek tussen onder andere leerproblemen, taalstoornissen en intellectuele disfuncties. Daarnaast biedt deze vorm van diagnostiek concrete handvatten om begeleidingen en behandelingen aan te passen aan de individuele sterktes en zwaktes vanuit het CHC profiel van de patiënt (bijvoorbeeld Fiorello & Primerano, 2005; Niileksela & Reynolds, 2014; Proctor, 2012; Uzieblo, Winter, Vanderfaeillie, Rossi, & Magez, 2012). Wil men zoveel mogelijk BCV's in kaart brengen, dan is men momenteel aangewezen op een crossbatterijbenadering, waarbij aan een bepaalde intelligentietest subtesten uit andere kwaliteitsvolle testen worden toegevoegd (zie bijvoorbeeld Flanagan et al., 2013). De Woodcock-Johnson III Tests of Cognitive and Achievement Abilities (WJ III; Woodcock, McGrew, & Mather, 2001), bij de ontwikkeling waarvan het CHC-model gehanteerd werd, biedt in de Angelsaksische landen momenteel het enige alternatief voor de crossbatterijbenadering. De eerste Vlaamse intelligentietest volgens het CHC-model is de Cognitieve Vaardigheidstest volgens het CHC-model (CoVaT-CHC, Magez et al., in press) waarvan de eerste versie binnenkort verschijnt. Deze basisversie zal nog niet alle BCV's dekken; de verwachting is dat dit in de volgende versies in grotere mate het geval zal zijn.

Voor het inschatten van een intellectuele disfunctie is het niet alleen van belang om een multidimensionele benadering van intelligentie toe te passen, maar ook om de klemtoon te leggen op het adaptief (dis)functioneren (American Assocation on Intellectual and Developmental Disabilities, 2015), zoals ook in de laatste editie van de DSM wordt beklemtoond (American Psychiatric Association, 2013). Onder adaptief functioneren verstaat men doorgaans drie categorieën van vaardigheden: (1) de conceptuele vaardigheden, waaronder taalvaardigheid, geletterdheid, en een begrip van geld, tijd en numerieke concepten; (2) de sociale vaardigheden, waaronder empathie, communicatievaardigheden en zelfregulatie, en (3) de praktische vaardigheden, waaronder geldbeheer, zelfzorg en organisatie van werk en school. De inschatting van het adaptief functioneren dient dus een prominentere rol te krijgen binnen de diagnostiek van het intellectueel functioneren, aangezien adaptief functioneren de beste indicator zou zijn voor de bepaling van de aard en de intensiteit van de begeleiding. Meer nog, volgens de criteria van de DSM-5 dient de ernst van de intellectuele disfunctie bepaald te worden aan de hand van het functioneren op de drie voorgenoemde domeinen. Hierbij dient er wel een duidelijke link te zijn tussen het intellectueel en adaptief disfunctioneren. Hoe men dit naar de testdiagnostiek kan vertalen, blijft nog de vraag (Greenspan & Woods, 2014). In Vlaanderen stuit men op een

bijkomend prangend probleem: voornamelijk voor de volwassen doelgroep ontbreekt het tot op heden aan kwaliteitsvolle instrumenten voor de meting van het adaptief functioneren die inhoudelijk alle relevante subdomeinen van het adaptief functioneren meten, die aan de psychometrische vereisten van validiteit en betrouwbaarheid voldoen, en die voor de Nederlandstalige algemene, laat staan de forensische, populatie genormeerd zijn (Vlaams Agentschap voor Personen met een Handicap, 2011-2012). Daarenboven verblijven forensische cliënten vaak in een gesloten, beveiligde omgeving, wat de evaluatie van het adaptief functioneren sterk bemoeilijkt. Het is immers binnen deze omgeving niet mogelijk om alle relevante gedragingen te observeren en/of derden (bijvoorbeeld de ouders) als informant te gebruiken, wat wel een vereiste is voor een betrouwbare en valide inschatting van het sociaal aanpassingsvermogen.

CONCLUSIE

Het belang van een kwaliteitsvolle meting van het intellectueel (dis)functioneren kan niet voldoende onderstreept worden. Dit is zo binnen de forensische werkcontext, maar geldt zonder twijfel ook in andere contexten. De beperkingen van intelligentietesten in rekenschap nemend, is een kritische houding ten aanzien van intelligentietesten en -scores dan ook onontbeerlijk. Dit is echter geen pleidooi om het kind met het badwater weg te gooien; de huidige intelligentietesten blijven een belangrijke tool om op objectieve wijze het intellectueel functioneren in kaart te brengen. Deze kritische houding dient zich bij voorkeur te vertalen in onder meer genuanceerde conclusies die men op basis van de intelligentiemeting trekt, en een voortdurend streven naar de verdere optimalisering van deze diagnostiek door onder meer het intellectueel functioneren op multidimensionele wijze te benaderen.

NOTEN

[1] In DSM-5 werd de term *verstandelijke beperking* of *mentale retardatie* vervangen door "een tekort in het intellectueel functioneren of een intellectuele disfunctie" (*Intellectual Disability*). In dit artikel zal daarom geopteerd worden om deze nieuwe benaming te hanteren.

[2] De WAIS-III is de derde herziene versie van de WAIS.

REFERENTIES

Alfonso, V.C., Flanagan, D.P., & Radwan, W. (2005). The impact of Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 185 – 202). New York, NY: Guilford Press.

American Assocation on Intellectual and Developmental Disabilities. (2015). Defintition of Intellectual Disability. Retrieved from http://aaidd.org/intellectual-disability/definition/faqs-on-intellectual-disability#.VRv8EbpsCSA

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Appelbaum, P. S. (2014). Hall v. Florida: Defining intellectual disability in the shadow of the death penalty. *Psychiatric Services, 65*(10), 1186-1188. doi: 10.1176/appi.ps.651004

Bellisle, F. (2004). Effects of diet on behaviour and cognition in children. *The British Journal of Nutrition, 92* (Suppl 2), S227-232.

Casey, P., & Keilitz, I. (1990). Estimating the prevalence of learning disabled and mentally retarded juvenile offenders: A meta-analysis. In P. E. Leone (Ed.), *Understanding roubled and troubling youth* (pp. 82-101). Newbury Park, CA: Sage.

Cosyns, P. (2005). *Werkgroep forensisch psychiatrisch zorgcircuit: Syntheseverslag mei 2005*. Brussel: Federale Overheidsdienst Justitie.

DeYoung, C. G., Peterson, J. B., Seguin, J. R., Mejia, J. M., Pihl, R. O., Beitchman, J. H., . . . Palmour, R. M. (2006). The dopamine D4 receptor gene and moderation of the association between externalizing behavior and IQ. *Archives of General Psychiatry, 63*(12), 1410-1416. doi: 10.1001/archpsyc.63.12.1410

Nuovo, A. G. Di, Nuovo, S. Di, & Buono, S. (2012). Intelligent quotient estimation of mental retarded people from different psychometric instruments using artificial neural networks. *Artificial Intelligence in Medicine, 54*(2), 135-145. doi: 10.1016/j.artmed.2011.11.002

Dias, S., Ware, R. S., Kinner, S. A., & Lennox, N. G. (2013). Co-occurring mental disorder and intellectual disability in a large sample of Australian prisoners. *The Australian and New Zealand Journal of Psychiatry, 47*(10), 938-944. doi: 10.1177/0004867413492220

Estevis, E., Basso, M. R., & Combs, D. (2012). Effects of practice on the Wechsler Adult Intelligence Scale-IV across 3- and 6-month intervals. *The Clinical Neuropsychologist*, *26*(2), 239–54. doi: 10.1080/13854046.2012.659219

Evans, I. M. (1991). Testing and diagnosis: A review and evaluation. In L. H. Meyer, C. A. Peck, & L. Brown (Eds.), *Critical Issues in the Lives of People with Severe Disabilities*. Baltimore, MD: Brookes Publishing Co.

Fiorello, C. A., & Primerano, D. (2005). Research into practice: Cattell-Horn-Carroll cognitive assessment in practice: Eligibility and program development issues. *Psychology in the Schools, 42*, 525-536.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). *Essentials of Cross-Battery Assessment* (3th ed.). New Jersey, NJ: John Wiley & Sons, Inc.

Floyd, R. G., Clark, M. H., & Shadish, W. R. (2008). The exchangeability of IQs: Implications for professional psychology. *Professional Psychology: Research and Practice, 39*, 414-423. doi: 10.1037/0735-7028.39.4.414

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 101*, 171-191.

Gimpel, G. A., Collett, B. R., Veeder, M. A., Gifford, J. A., Sneddon, P., Bushman, B., . . . Odell, J. D. (2005). Effects of stimulant medication on cognitive performance of children with ADHD. *Clinical Pediatrics 44*(5), 405-411. doi: 10.1177/000992280504400504

Greenspan, S., & Woods, G. W. (2014). Intellectual disability as a disorder of reasoning and judgement: the gradual move away from intelligence quotient-ceilings. *Current Opinion in Psychiatry, 27*(2), 110-116. doi: 10.1097/YCO.0000000000000037

Habets, P., Jeandarme, I., Uzieblo, K., Oei, T. I., & Bogaerts, S. (2014). Intelligence is in the eye of the beholder: Investigating repeated IQ measurements in forensic psychiatry. *Journal of Applied Research in Intellectual Disabilities*, *28*(3), 182-192. doi: 10.1111/jar.12120

Heaton, R. K., Gladsjo, J. A., Palmer, B. W., Kuck, J., Marcotte, T. D., & Jeste, D. V. (2001). Stability and course of neuropsychological deficits in schizophrenia. *Archives of General Psychiatry, 58*(1), 24-32.

Herrnstein, R. J., & Murray, C. (1994). *The Bell curve: Intelligence and class structure in American life*. New York, NY: The Free Press.

Jones, J. (2007). Persons with intellectual disabilities in the criminal justice system: Review of issues. *International Journal of Offender Therapy and Comparative Criminology, 51*(6), 723-733. doi: 10.1177/0306624X07299343

Kassin, S. M. (2008). The Psychology of confessions. *Annual Review of Law and Social Science, 4*(1), 193-217. doi: 10.1146/annurev.lawsocsci.4.110707.172410

Keith, T. Z., & Reynolds, M. R. (2010). Cattell–Horn–Carroll abilities and cognitive tests: What we've learned from 20 years of research *Psychology in the Schools, 47*(7), 635-650. doi: 10.1002/pits.20496

Koenen, K. C., Caspi, A., Moffitt, T. E., Rijsdijk, F., & Taylor, A. (2006). Genetic influences on the overlap between low IQ and antisocial behavior in young children. *Journal of Abnormal Psychology, 115*(4), 787-797. doi: 10.1037/0021-843X.115.4.787

Kooreman, A., & Luteijn, F. (1987). *The short format of the Groningen Intelligence Test*. Lisse: Swets & Zeitlinger.

Koyama, K. I., Asakawa, A., Nakahara, T., Amitani, H., Amitani, M., Saito, M., . . . Inui, A. (2012). Intelligence quotient and cognitive functions in severe restricting-type anorexia nervosa before and after weight gain. *Nutrition, 28*(11-12), 1132-1136. doi: 10.1016/j.nut.2012.03.003

Lahey, B. B., D'Onofrio, B. M., Hulle, C. A. Van, & Rathouz, P. J. (2014). Prospective association of childhood receptive vocabulary and conduct problems with self-reported adolescent delinquency: Tests of mediation and moderation in sibling-comparison analyses. *Journal of Abnormal Child Psychology, 42*(8), 1341-1351. doi: 10.1007/s10802-014-9873-x

Levin, R., Dor-Abarbanel, A. E., Edelman, S., Durrant, A. R., Hashimoto, K., Javitt, D. C., & Heresco-Levy, U. (2015). Behavioral and cognitive effects of the N-methyl-D-aspartate receptor co-agonist D-serine in healthy humans: Initial findings. *Journal of Psychiatric Research, 61*, 188-195. doi: 10.1016/j.jpsychires.2014.12.007

Lindsay, W. R. (2011). People with intellectual disability who offend or are involved with the criminal justice system. *Current Opinion in Psychiatry, 24*(5), 377-381. doi: 10.1097/YCO.0b013e3283479dc9

Lindsay, W. R., Hastings, R. P., & Beail, N. (2013). Why do some people with intellectual disability engage in offending behaviour and what can we do about it? *Journal of Applied Research in Intellectual Disabilities, 26*(5), 351-356. doi: 10.1111/jar.12042

Loeber, R., Menting, B., Lynam, D. R., Moffitt, T. E., Stouthamer-Loeber, M., Stallings, R., . . . Pardini, D. (2012). Findings from the Pittsburgh Youth Study: Cognitive impulsivity and intelligence as predictors of the age-crime curve. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(11), 1136-1149. doi: 10.1016/j.jaac.2012.08.019

MacLean, H., McKenzie, K., Kidd, G., Murray, A. L., & Schwannauer, M. (2011). Measurement invariance in the assessment of people with an intellectual disability. *Research in Developmental Disabilities, 32*(3), 1081-1085. doi: 10.1016/j.ridd.2011.01.022

Maes, B., Goethals, J., & Verlinden, S. (2009). *Personen met een verstandelijke beperking onderhevig aan een interneringsmaatregel*. Leuven: Steunpunt Welzijn, Volksgezondheid en Gezin.

Magez, W., Tierens, M., Huynegem, J. Van, Parijs, K. Van, Decaluwé, V., & Bos, A. (In press). *CoVaT-CHC Basisversie. Cognitieve Vaardigheidstest volgens het CHC-model*. Antwerpen: Thomas More.

Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T., & Saxena, S. (2011). Prevalence of intellectual disability: A meta-analysis of population-based studies. *Research in Developmental Disabilities, 32*(2), 419-436. doi: 10.1016/j.ridd.2010.12.018

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*(1), 1-10. doi: 10.1016/j.intell.2008.08.004

Mears, D. P., & Cochran, J. C. (2013). What is the effect of IQ on offending. *Criminal Justice and Behavior, 40*(11), 1280-1300. doi: 10.1177/0093854813485736

Niileksela, C. R., & Reynolds, M. R. (2014). Global, broad, or specific cognitive differences? Using a MIMIC model to examine differences in CHC abilities in children with learning disabilities. *Journal of Learning Disabilities, 47*(3), 224-236. doi: 10.1177/0022219412453172

Portnoy, J., Chen, F. R., & Raine, A. (2013). Biological protective factors for antisocial and criminal behavior. *Journal of Criminal Justice, 41*(5), 292-299. doi: 10.1016/j.jcrimjus.2013.06.018

Proctor, B. (2012). Relationships between Cattell-Horn-Carroll (CHC) cognitive abilities and math achievement within a sample of college students with learning disabilities. *Journal of Learning Disabilities, 45*(3), 278-287. doi: 10.1177/0022219410392049

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.

Reynolds, M. R., Ingram, P. B., Seeley, J. S., & Newby, K. D. (2013). Investigating the structure and invariance of the Wechsler Adult Intelligence Scales, Fourth Edition in a sample of adults with intellectual disabilities. *Research in DevelopmentalDdisabilities, 34*(10), 3235-3245. doi: 10.1016/j.ridd.2013.06.029

Schalock, R. L., Borthwick-Duffy, S. A., Bradley, V. J., Buntinx, W. H. E., Coulter, D. L., Craig, E. M., . . . Yeager, M. H. (2010). *Intellectual Disability: Definition, Classification, and Systems of Supports* (11th ed.). Washington, D.C: American Association on Intellectual and Developmental Disabilities.

Staff, R. T., Hogan, M. J., & Whalley, L. J. (2014). Aging trajectories of fluid intelligence in late life: The influence of age, practice and childhood IQ on Raven's Progressive Matrices. *Intelligence*, *47*, 194–201. doi:10.1016/j.intell.2014.09.013

Silverman, W., Miezejeski, C., Ryan, R., Zigman, W., Krinsky-McHale, S., & Urv, T. (2010). Stanford-Binet & WAIS IQ Differences and Their Implications for Adults with Intellectual Disability (aka Mental Retardation). *Intelligence, 38*(2), 242-248. doi: 10.1016/j.intell.2009.12.005

Smithers, L. G., Golley, R. K., Mittinty, M. N., Brazionis, L., Northstone, K., Emmett, P., & Lynch, J. W. (2012). Dietary patterns at 6, 15 and 24 months of age are associated with IQ at 8 years of age. *European Journal of Epidemiology, 27*(7), 525-535. doi: 10.1007/s10654-012-9715-5

Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The Flynn effect: A meta-analysis. *Psychological Bulletin, 140*(5), 1332-1360. doi: 10.1037/a0037173

Toorn, B. Van, & Bon, C. (2011). De onbetrouwbaarheid van de IQ-meting [The unreliability of IQ-tests]. *De Psycholoog, 45*, 4-9.

Winkel, R. van, Myin-Germeys, I., Delespaul, P., Peuskens, J., Hert, M. De, & Os, J. van (2006). Premorbid IQ as a predictor for the course of IQ in first onset patients with schizophrenia: a 10-year follow-up study. *Schizophrenia Research, 88*(1-3), 47-54. doi: 10.1016/j.schres.2006.06.033

Uzieblo, K., Winter, J., Vanderfaeillie, J., Rossi, G., & Magez, W. (2012). Intelligent diagnosing of intellectual disabilities in offenders: Food for thought. *Behavioral Sciences & the Law, 30*(1), 28-48. doi: 10.1002/bsl.1990

Vlaams Agentschap voor Personen met een Handicap (VAPH). (2011-2012). Classificerende diagnostische protocollen. Retrieved from www.vaph.be/vlafo/download/nl/8229547/bestand

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale WAIS*. New York, NY: Psychological Corporation.

Wechsler, D. (1970). *WAIS Nederlandstalige bewerking. Technische Handleiding*. Amsterdam: Swets & Zeitlinger.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale, Third ed. administration and scoring manual.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (2005). W*AIS-III Nederlandstalige bewerking. Technische Handleiding*. Lisse: Swets & Zeitlinger.

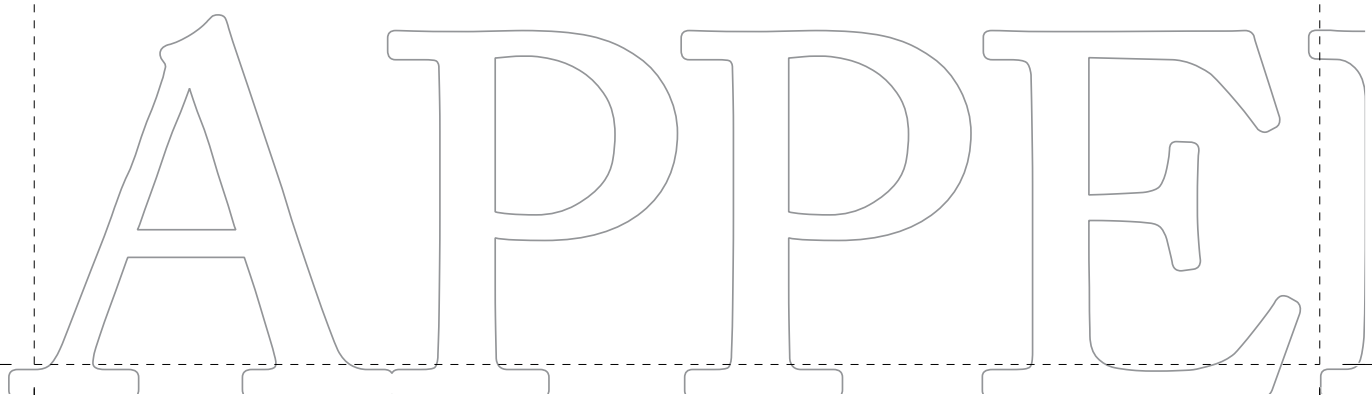
Wechsler, D. (2012). *WAIS-IV Nederlandstalige bewerking. Technische Handleiding.* Lisse: Swets & Zeitlinger.

Whitaker, S. (2008). Intellectual disability: A concept in need of revision? *The British Journal of Developmental Disabilities, 54*, 3-9.

Whitaker, S. (2010). Error in the estimation of intellectual ability in the low range using the WISC-IV and WAIS-III. *Personality and Individual Differences, 48*(5), 517-521. doi: 10.1016/j.paid.2009.11.017

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing.

Zampera, E. (1999). Intellectual performance of chronic schizophrenic patients. *Collegium Antropologicum, 23*(2), 597-602.



Validity and reliability of the VRAG in a forensic psychiatric medium security population in Flanders

Van Heesch, B., Jeandarme, I., Pouls, C., & Vervaeke, G.

Psychology, Crime & Law, 22(6), 530-537

APPENDIX

ABSTRACT

Since the development of the Violence Risk Appraisal Guide (VRAG), one of the most widely used actuarial risk assessment instruments, numerous replication studies have shown its usefulness in predicting (violent) recidivism among various offender populations. It is not clear, however, whether these findings can be generalized to forensic psychiatric patients admitted to a medium security forensic psychiatric unit in the Flemish part of Belgium. Therefore, the main aim of this study was to test the predictive validity and reliability of the VRAG in a sample of 191 Flemish medium security patients. The mean follow-up period was 2.4 years. Contrary to the expectations, the VRAG was unable to significantly predict any kind of outcome. Possible explanations are discussed and further research with the VRAG in Flanders, and its recently revised version, is recommended.

INTRODUCTION

The Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice, & Cormier, 1998, 2006) is an actuarial instrument developed in Canada for the prediction of violent recidivism (Harris, Rice, & Quinsey, 1993). Since its development, more than 70 replication studies have been conducted worldwide (Penetanguishene Research Department, 2015; Rice, Harris, & Lang, 2013). In general, most studies confirm the moderate to large predictive accuracy of this instrument for various offender populations and various outcomes (e.g., Pham, Ducro, Marghem, & Réveillère, 2005; Rice et al., 2013; Sjöstedt & Långström, 2002). In a recent meta-analysis of Singh, Grann, and Fazel (2011), the median area under the curve (AUC) for the VRAG was .74. Additional performance indicators showed that 66% of the individuals judged to be at high risk actually recidivated (positive predictive value, PPV). The negative predictive value (NPV) indicated that 74% of the participants judged to be at low risk did not go on to recidivate. Although the VRAG successfully predicted recidivism in several international studies, it is not clear whether these results can be generalized to forensic psychiatric patients admitted to a medium security unit in the Flemish part of Belgium. Therefore, the aim of the current study was to test the reliability and predictive accuracy of the VRAG for this population. So far, only one study has been conducted using the Dutch version of the VRAG (Pouls, Jeandarme, & Habets, 2014). The sample in this prospective research consisted of 52 male offenders with an intellectual disability. The interrater reliability was good and the internal consistency acceptable, but the VRAG was unable to significantly predict any kind of institutional incident or violence. It must be noted that the small study sample and low base rates could be an explanatory factor for this low predictive accuracy.

METHOD

Participants

The sample consisted of 244 male forensic psychiatric patients who were admitted to one of the three forensic medium security units in Flanders between 2001 and 2010. In 17 (6.97%) cases, no valid VRAG score could be obtained due to incomplete or unreliable file information. Participants who were not discharged before the end of the study period were excluded as well (n = 36).

Therefore, 191 subjects were included in the final sample. The mean age in this sample was 37.15 years (*SD* = 10.99, range 19–70). Most of the participants (88.42%) had Belgian nationality. The mean WAIS-III IQ score (Wechsler, 2000) was 81.46 (*SD* = 16.48, range 48–138). According to the Diagnostic and Statistical Manual of mental disorders-IV-Text Revision (DSM-IV-TR; American Psychiatric Association, 2000), the most common mental disorders among the study sample were personality disorders (75.39%), substance-related disorders (60.21%), psychotic disorders (45.03%), and intellectual disability (13.61%). The majority (80.95%) had been admitted to a psychiatric hospital prior to the index offense. Also, most of the participants (92.15%) had been convicted and/or interned for violent offenses. The index offense was in 76.96% of the cases a (sexual) violent offense, in 17.28% a property offense, in 2.62% a drug offense, in 2.09% a hands-off sex offense, and in 1.05% vandalism.

Instruments

The VRAG consists of 12 items with a total score ranging between −26 and +38, higher scores indicating a higher probability of a violent outcome. For 46.07% of the participants, one or several items were dropped due to insufficient (reliable) file information. For example, the psychopathy item was dropped in 12.04% (*n* = 23) of the cases. The Childhood and Adolescent Taxon Scale (CATS) (Jeandarme, Pouls, & Peters, 2012; Quinsey et al., 2006), which counts as a valid alternative for the Psychopathy Checklist-Revised (Hare, 2003) in scoring the VRAG, was used in half of the remaining cases (48.81%, *n* = 82). Scores on the VRAG varied from −26 to 35, with a mean score of 7.14 (*SD* = 10.87). Only valid total scores (no more than four missing items) were included in the final sample (supra). For the current study purpose, the recently translated Dutch versions (Jeandarme et al., 2012) of the VRAG and CATS were used.

Procedure

Two raters of the Knowledge Center Forensic Psychiatric Care were trained by one of the authors of the VRAG and were supervising the other rater. Ratings were based on the file information. This research project was formally approved by the Ethics committee of Antwerp University Hospital.

Time at risk and outcome measures.

The mean time at risk, defined as the period after release from the medium security unit until the end of 2010 or earlier in case of death, was 2.44 years (*SD* = 796.75 days, range 1–3165 days). After their release, offenders were referred to a less secure setting, including, for example, community, non-forensic psychiatric hospitals, assisted living facilities, and (closed) long-term care facilities. Two outcome measures were taken into account, namely *violent recidivism* and *violent offense*. Violent recidivism was defined as a new conviction or internment[1] for a violent (sexual) offense based on the criminal record of the participants. As all of the participants were already under supervision of the judicial authorities, violent behavior did not always result in a reconviction. Therefore, all violent incidents reported to the judicial authorities were added to official reconviction rates to account for a second outcome measure. The term *violent offense* will be further used to refer to this second outcome measure.

Data analyses

Data were analyzed with IBM SPSS for Windows version 22. All tests were two tailed with significance level .05 (5%). Although reliability is a necessary condition for validity (Quinsey et al., 2006), controversy exists concerning how to measure the reliability of the VRAG and other actuarial risk assessment instruments. Some authors limit this psychometric property by merely calculating the interrater reliability (e.g., Quinsey et al., 2006), while others also calculate Cronbach's α (e.g., Loza, Villeneuve, & Loza-Fanous, 2002; Vitacco, Gonsalves, Tomony, Smith, & Lishner, 2012). Despite Cronbach's α being a commonly used measure to calculate the internal consistency of risk assessment instruments, it could be argued whether this statistic is an appropriate standard for evaluating the reliability of such instruments. An important consideration in this respect is whether the VRAG uses a reflective or formative measurement model (Jaspaert, 2015). In reflective measurement models, items are indicators of the measured construct and are *caused* by this construct. Hence, items overlap in meaning and are supposed to have high intercorrelations (Baxter, 2009). As such, Cronbach's α can be meaningfully applied (Coltman, Devinney, Midgley, & Venaik, 2008). In formative measurement models, however, items are independent causes of the measured construct and are not expected to correlate highly with each other (Baxter, 2009). Consequently, Cronbach's α may not be the most appropriate measure

for such models (Coltman et al., 2008). Accordingly, the VRAG can be viewed as a formative measure, since the total score is formed by the individual items, and not the other way around. Therefore, change in one item is not necessarily associated with changes in other items (Jaspaert, 2015). In this respect, the interrater reliability and item-total correlations, indicating the correlation between each individual item and the total score of the VRAG, can be seen as more adequate reliability indicators than Cronbach's α. For the current study purpose, Cronbach's α, the interrater reliability, as well as item-total correlations were calculated. The interrater reliability of the VRAG was examined in 27 participants through intraclass correlation coefficients (ICC; absolute agreement, two-way random, single measurement).

Predictive validity was analyzed using seven performance indicators. A global effect size was calculated through the receiver operating characteristic (ROC; Mossman, 1994; Swets, 1992), with AUC values indicating the probability of a randomly selected recidivist having a higher risk classification than a randomly selected non-recidivist. AUCs can range from 0 to 1, with an AUC of 1 representing a totally accurate prediction and an AUC of .50 a prediction equal to chance (Rice & Harris, 1995). Using the information of a 2 × 2 contingency table, sensitivity (percentage of recidivists who were judged to be at high risk), specificity (percentage of non-recidivists who were judged to be at low risk), PPV (percentage of participants judged to be at high risk who went on to recidivate), NPV (percentage of low-risk individuals who did not go on to recidivate), number needed to detain (NND; number of participants judged to be at high risk who need to be detained to prevent a single incident or offense), and number safely discharged (NSD; number of participants judged to be at low risk who could be discharged prior to a single incident or offense) were calculated to further evaluate predictive accuracy (Singh, 2013). These performance indicators provide information about how accurate a tool identifies high-risk (*rule in*; PPV and NND) and low-risk (*rule out*; NPV and NSD) individuals. In order to calculate these measures, patients in risk category 7 or higher (scores of 14 or higher) were considered high risk.

RESULTS

Base rates

During time at risk, 4.71% (*n* = 9) of the participants was reconvicted for a violent offense. Adding all registered violent offenses to the official recidivism data increased the base rate to 19.90% (*n* = 38).

Reliability of the VRAG

The ICC was high (.91), indicating a good interrater reliability. Cronbach's α was .63. As can be seen from Table 1, item-total correlation tests showed that eight items did not correlate very well (< .3) with the total score of the VRAG.

Table 1. Item-total correlations of the VRAG

Items VRAG	Item-total correlation
Lived with both biological parents to age 16	.40
Elementary school maladjustment	.57
History of alcohol problems	.26
Marital status	.13
Criminal history score for nonviolent offenses	.32
Failure on prior conditional release	.22
Age at index offense	.27
Victim injury	.14
Any female victim	.18
Meets DSM-III criteria for any personality disorder	.19
Meets DSM-III criteria for schizophrenia	.04
Hare Psychopathy Checklist-Revised score	.59

Predictive validity of the VRAG

Table 2 shows the performance indicators for both outcome measures. In general, the VRAG failed to significantly predict any kind of outcome, with AUCs around chance level. Up to one-third of the recidivists were judged to be at high risk for reoffending (sensitivity = 21.05–33.33%). About two-thirds of the non-recidivists was classified as low risk (specificity = 69.93–71.98%). The

VRAG was not useful in prospectively predicting who was going to recidivate (PPV = 5.56–14.81%), but it did identify low-risk patients with high accuracy (NPV = 78.10–95.62%).

Table 2. Predictive validity of the VRAG

	Violent recidivism	Violent offence
AUC total score (CI)	.57 (.39–.74)	.46 (.37–.56)
AUC bin (CI)	.55 (.37–.72)	.46 (.36–.55)
Sensitivity	33.33	21.05
Specificity	71.98	69.93
PPV	5.56	14.81
NPV	95.62	78.10
NND	18	7
NSD	22	4

Note. AUC = area under the curve; PPV = positive predictive value; NPV = negative predictive value; NND = number needed to detain; NSD = number safely discharged; CI = confidence interval.
* $p < .05$.

DISCUSSION

The main aim of this study was to test the reliability and predictive validity of the VRAG for offenders admitted to a medium security unit in Flanders. Interrater reliability was high but internal consistency was questionable, with over half of the items that did not correlate well with the total score. Furthermore, the VRAG was only accurate in identifying low risk individuals.

The interrater reliability of the VRAG was good (ICC = .91) and consistent with various international studies (e.g., Harris, Rice, & Cormier, 2002; Pouls et al., 2014). The internal consistency was acceptable for research purposes (α = .63), but item-total correlation tests showed that more than half of the individual items did not correlate very well with the total score of the VRAG. These reliability results are consistent with various international studies (e.g., Loza et al., 2002: α = .72; Vitacco et al., 2012: α = .67) and the first study of the Dutch version of the VRAG in offender with intellectual disabilities, where an ICC of .90 and an α of .61 was found (Pouls et al., 2014). As stated above, the modest α coefficient has to be qualified by considering the VRAG as a formative measurement model, since change in an individual item is not

necessarily associated with changes in other items (Jaspaert, 2015). The low item-total correlations, however, are more problematic, but with still an interrater reliability well above average.

The VRAG failed, contrary to the expectations, to discriminate violent recidivists from non-violent recidivists, with AUCs no better than chance. In addition, the tool was not able to predict who was going to recidivate. On the other hand, NPV and NSD were rather high, indicating that the VRAG is effective in making rule out decisions. These findings are in sharp contrast with most (international) studies, where AUCs between .60 and .80 were found in forensic psychiatric patients (e.g., Harris et al., 2002; Ho, Thomson, & Darjee, 2009; Pham et al., 2005). Several limitations of this study may have caused this low predictive validity. First, the results could be influenced by the retrospective design of the study. Furthermore, information about the offender's youth was often minimal, totally lacking and/or solely based on self-report. Second, the base rate for violent recidivism was low, resulting in a loss of statistical power. As low base rates were expected, all reported violent offenses were added in the current study. Although the base rate of all reported offenses was sufficient, AUCs remained below chance level. Furthermore, these findings are in contrast with the Belgian study of Pham et al. (2005), which showed high predictive values for violent recidivism (AUC = .82) with a base rate of 11.8%. Third, differences between the study sample and the sample in the original study (Harris et al., 1993) could have influenced the results. In the current study, only offenders admitted to a medium security unit were included, while the construction sample included both forensic psychiatric patients and inmates. However, other studies found significant effects in forensic psychiatric patients (e.g., Doyle & Dolan, 2006; Snowden, Gray, Taylor, & MacCulloch, 2007). Fourth, it must be noted that patients often still reside in a structured and/or (heavily) secured setting after their release from the medium security unit. Because not all patients are directly released into the community, it is possible that they, for example, stay in an assisted living facility or even a (closed) long-term care facility. The low rates of reoffending – despite the relatively high scores on the VRAG – could, therefore, be a consequence of successful risk management. Fifth, the follow-up period included variable time periods. Analyses with a minimal follow-up period of one year, however, did not change the results.

CONCLUSIONS AND RECOMMENDATIONS

In the current study, the VRAG failed to significantly predict violent re-offenses. This is remarkable in the light of extensive previous international research suggesting that the (predictive) validity of the VRAG is good. In addition, it was argued whether standards for measuring internal consistency reliability can be applied to an actuarial risk assessment instrument. Further research should shed more light on the validity of the VRAG in medium security samples in Belgium. More recently, Harris, Rice, Quinsey, and Cormier (2015) developed a revised and easier-to-score version of the VRAG: the Violence Risk Appraisal Guide-Revised (VRAG-R). The use of this revised instrument could possibly lead to better results. For now, clinicians should remain prudent for the blind use of any kind of risk assessment instrument that is not validated for a certain population and setting.

FOOTNOTES

[1] Henceforth simply referred to as *reconviction*.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of mental disorders* (Text Revision, 4th ed.). Washington, DC: American Psychiatric Association.

Baxter, R. (2009). Reflective and formative metrics of relationship value: A commentary essay. *Journal of Business Research, 62*(12), 1370-1377. doi: 10.1016/j.jbusres.2008.12.004

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research, 61*(12), 1250-1262. doi: 10.1016/j.jbusres.2008.01.013

Doyle, M., & Dolan, M. (2006). Predicting community violence from patients discharged from mental health services. *The British Journal of Psychiatry, 189*(6), 520-526. doi: 10.1192/bjp.bp.105.021204

Harris, G. T., Rice, M. E., & Cormier, C. A. (2002). Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior, 26*(4), 377-394.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior, 20*(4), 315-335. doi: 10.1177/0093854893020004001

Harris, G. T., Rice, M. E., Quinsey, V. L., & Cormier, C. A. (2015). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Ho, H., Thomson, L., & Darjee, R. (2009). Violence risk assessment: The use of the PCL-SV, HCR-20, and VRAG to predict violence in mentally disordered offenders discharged from a medium secure unit in Scotland. *The Journal of Forensic Psychiatry & Psychology, 20*(4), 523-541. doi: 10.1080/14789940802638358

Jaspaert, E. (2015). *A dyadic analysis of the role of preference disconfirmation in the explanation of intimate partner violence.* (Unpublished doctoral thesis). University of Leuven, Leuven, Belgium.

Jeandarme, I., Pouls, C., & Peters, M. J. V. (2012). *Violence Risk Appraisal Guide: Richtlijnen om het risico op (seksueel) gewelddadig gedrag te beoordelen*. Hasselt: Leën.

Loza, W., Villeneuve, D. B., & Loza-Fanous, A. (2002). Predictive validity of the Violence Risk Appraisal Guide: A tool for assessing violent offender's recidivism. *International Journal of Law and Psychiatry, 25*(1), 85-92.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting Clinical Psychology, 62*(4), 783-792.

Penetanguishene Research Department. (2015). *Replications of the Violence Risk Appraisal Guide or Sex Offender Risk Appraisal Guide in Assessing Violence Risk*. Retrieved from http://static1.squarespace.com/static/520a76a0e4b03ad27abae1e3/t/53fa22ffe4b017241945ee72/1408901887175/RagReps.pdf

Pham, T. H., Ducro, C., Marghem, B., & Réveillère, C. (2005). Évaluation du risque de récidive au sein d'une population de délinquants incarcérés ou internés en Belgique francophone [Prediction of recidivism among prison inmates and forensic patients in Belgium]. *Annales Médico-psychologiques, 163*(10), 842-845. doi: 10.1016/j.amp.2005.09.013

Pouls, C., Jeandarme, I., & Habets, P. (2014). Risicotaxatie bij daders met een verstandelijke beperking. Eerste toepassing Nederlandstalige VRAG [Risk assessment in offenders with intellectual disability: A first application of the Dutch VRAG]. *De Psycholoog, 1*(42-51).

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Rice, M. E., Harris, G. T., & Lang, C. (2013). Validation of and revision to the VRAG and SORAG: The Violence Risk Appraisal Guide-Revised (VRAG-R). *Psychological Assessment, 25*(3), 951-965. doi: 10.1037/a0032878

Singh, J. P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences & the Law, 31*(1), 8-22. doi: 10.1002/bsl.2052

Singh, J. P., Grann, M. n., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499-513. doi: 10.1016/j.cpr.2010.11.009

Sjöstedt, G., & Långström, N. (2002). Assessment of risk for criminal recidivism among rapists: A comparison of four different measures. *Psychology, Crime & Law, 8*(1), 25-40. doi: 10.1080/10683160208401807

Snowden, R. J., Gray, N. S., Taylor, J., & MacCulloch, M. J. (2007). Actuarial prediction of violent recidivism in mentally disordered offenders. *Psychological Medicine, 37*(11), 1539-1549. doi: 10.1017/s0033291707000876

Swets, J. A. (1992). The science of choosing the right decision threshold in high-stakes diagnostics. *American Psychologist, 47*(4), 522-532.

Vitacco, M. J., Gonsalves, V., Tomony, J., Smith, B. E. R., & Lishner, D. A. (2012). Can standardized measures of risk predict inpatient violence? Combining static and dynamic variables to improve accuracy. *Criminal Justice and Behavior, 39*(5), 589-606. doi: 10.1177/0093854812436786

Wechsler, D. (2000). *WAIS-III Nederlandstalige Bewerking. Technische Handleiding*. Lisse: Swets & Zeitlinger.

What started as the implementation of the first research project in a forensic psychiatric hospital in September, 2009, ended in a 5-year adventure in writing this. Looking back on those years, I would like to thank all the people who supported me during that period.

First of all, I want to thank John Vanacker who prompted this journey and the Public Psychiatric Care Center (OPZC) in Rekem and *Limburg Sterk Merk*, who provided necessary research funds for several studies. Without them, this project would not have been possible.

A very special thank you to my two co-workers at OPZC, Claudia Pouls and Petra Habets, for supporting me, helping me with research projects and answering all of my questions. Thanks for sharing your knowledge and thoughts with me. Thank you, Claudia, for being such a perfectionist, and the same to Petra for always staying calm!

Without the chairs' help from the different Commissions for the Protection of Society, this project would not have been possible. In particular, thank you Henri, for all your support and suggestions!

My deep gratitude is extended to all patients who participated in the research projects, the clinicians who took the initiative to start data registration in 2001 (Ludwig De Groeve, Dominiek Derore, Lies Mariën, Rudy Reusens, Greet Stans, Ali Van der Auwera, Raf Vermeiren, and Wenke Wouters), and all other staff-members from the three medium security units: OPZC, the psychiatric center Sint-Jan-Baptist in Zelzate, and the university psychiatric center Sint-Kamillus in Bierbeek. Thanks to the steering committee – Vanessa Baetens, Jan De Laender, Jan De Varé, Steven Degrauwe, Kevin Pesout, and Rudi Verelst – for close collaboration and the time you all invested in making these studies possible and answering all additional questions. Also, the Ministry of Justice was more than helpful, as were Yves Van Den Berge, Francis Van Mol and Luc Holsters.

I appreciated the constructive feedback I received from co-authors including Freya Vander Laenen, Ciska Wittouck, Katrien Hanoulle, Laura Vandenbosch and Liesbeth Bruckers as well as from the anonymous reviewers of the different articles. Also, special thanks to the members of the doctoral guidance committee.

This work would not have been possible without enthusiastic supervisors. I am very thankful to prof. dr. Karel Oei, for his continuous support, encouragement, and the trust both he

and prof. dr. Groenhuijsen had in me at the beginning of this process. Karel and Marc, I would never have started this adventure without your faith in this project.

Special thanks and appreciation to my third supervisor, prof. dr. Stefan Bogaerts for his close collaboration, never-ending advice and encouragement. His ongoing support and constructive remarks always offered clear views on the direction of my work. Stefan, I could not have finished this adventure without you!

Lastly, I would like to express my appreciation to my friends, colleagues, and family for being there and encouraging me to continue, although it did take up most of my spare time in the past few years. I have saved the last word of acknowledgment for my husband and children for being by my side while doing research and writing this dissertation. Luc, Sam, and Max, from now on, I will have time to do fun things, but then again, *nothing really ends*.

Thanks with all my heart!

Ingeborg Jeandarme was born on February 23, 1963 in Sint-Truiden, Belgium. She studied medicine at the Universiteit Diepenbeek, and subsequently at the Katholieke Universiteit Leuven (KUL), where she got her diploma in July, 1988. From 1988 through 1993, she followed internship training in psychiatry under the supervision of Prof. F. Baro in Bierbeek. She rotated through a variety of departments at the psychiatric center Ziekeren (Sint-Truiden), Ziekenhuis Zuid-Oost Limburg (ZOL, Genk), and university psychiatric center Sint-Kamillus (Bierbeek). She also had one year of neurological training in ZOL and one year of forensic psychiatry training in the Central Prison in Leuven. During the last years of internship, she studied criminology at the KUL, where she obtained her diploma in 1992. As a psychiatrist, she took a three-year specific course in psychoanalytical psychotherapy until 1995. Further, she took courses in forensic psychiatry at the Katholieke Universiteit Nijmegen (1995), and followed the interuniversity course in forensic psychiatry/psychology at the KUL.

After becoming a certified psychiatrist, she started working in the psychiatric center Grote Beek (Eindhoven) from 1993 to 2009. She worked primarily as a psychiatrist on several wards for psychotic patients. During the last years, she was involved in risk education and took part in several working groups for the Task Force of the Expertisecentrum Forensische Psychiatrie (EFP, Utrecht). During her work in the Grote Beek, she was posted to Venray (Rooyse Wissel, Review Board for leaves, 2006 to 2009), and Maastricht (project group Penitentiair Psychiatrisch Ziekenhuis Overmaze, 2007 to 2008). Besides her work in the Netherlands, she worked for the justice department in Belgium from 1993 to 2008. Here, she was responsible for the treatment of internees within the prison system and became involved in the education of prison psychologists by giving workshops on risk assessment and psychopathy. From 2006 to 2009, she was a member of the Nationale Penitentiaire Gezondheidsraad. At the same time, she was a consulting psychiatrist in the Centrum Geestelijke Gezondheidszorg (Leuven), where she was responsible for the forensic psychiatric patients (1992 to 2002), as a therapist and expert assessing criminal responsibility evaluations. From 2002 to 2015, she had a small private practice at her home. From 2007 to 2014, she worked a few hours each week as a consulting psychiatrist for a new forensic project for intellectually disabled offenders (Limes, Sint-Truiden). From 2005 to 2013, she co-chaired the section on forensic psychiatry of the Vlaamse Vereniging voor Psychiatrie (VVP). The above part-time psychiatric experiences were within forensic settings, although she worked full-

time as a general psychiatrist in Ziekeren since 1998, where she was responsible for crisis and suicidal patients in the acute unit.

Currently she continues working full-time as a psychiatrist in Asster (formerly Ziekeren), where she is responsible for the treatment of patients with illegal substance misuse and paraphilic disorders. Since 2009, she works as a part-time coordinator at the Knowledge Center Forensic Psychiatric Care (KeFor), OPZC Rekem. Her main research topics include forensic psychiatry, internment, risk assessment, and recidivism. She is chair of the Lokale Kwaliteitsgroep (LOK) in Sint-Truiden and is an acting member of the Commission of the Protection of Society in Leuven. Since 2014, she teaches Gerechtelijke Geestelijke Gezondheidszorg together with prof. dr. Aertsen at the KUL Law Department.

Her doctoral research officially started in 2011 at the Tilburg Law School in the Netherlands. She gave lectures on parts of the dissertation, with other research at KeFor. Since 1997, she has been married to Luc. Her daughter Sam was born in 1998 and her son Max was born in 2001.

PEER REVIEWED JOURNALS

Jeandarme, I., Pouls, C., De Laender, J., Oei, T. I., & Bogaerts, S. (in press). Field validity of the HCR-20 in forensic medium security units in Flanders. *Psychology, Crime & Law*.

Jeandarme, I., Edens, J., Habets, P., Bruckers, L., Oei, T. I., & Bogaerts, S. (in press). PCL-R field validity in prison and hospital settings. *Law and Human Behavior*.

Pouls, C., & **Jeandarme, I.** (in press). Predicting institutional aggression in offenders with intellectual disabilities with the Violence Risk Appraisal Guide. *Journal of Applied Research in Intellectual Disabilities.*

Jeandarme, I., Wittouck, C., Laenen, F. Vander, Pouls, C., Oei, T. I., & Bogaerts, S. (2016). Risk factors associated with inpatient violence during medium security treatment. *Journal of Interpersonal Violence*. doi: 10.1177/0886260516670884

Jeandarme, I. (2016). Forensische psychiatrie à la flamande. *Orde van de* Dag, *74*, 32-40.

Jeandarme, I., Pouls, C., Hanoulle, K., Oei, T. I., & Bogaerts, S. (2016). Klinische implicaties van afschaffing van de ministeriële internering in België: Onderzoek naar profiel van Vlaamse geïnterneerden-veroordeelden [Abolition of the internment of convicted prisoners in Belgium: Clinical implications]. *Tijdschrift voor Psychiatrie*, *58*(8), 574-582.

Jeandarme, I., Habets, P., Oei, T. I., & Bogaerts, S. (2016). Reconviction and revocation rates in Flanders after medium security treatment. *International Journal of Law and Psychiatry*, *47*, 45-52. doi: 10.1016/j.ijlp.2016.02.033

Heesch, B. van, **Jeandarme, I.**, Pouls, C., & Vervaeke, G. (2016). Validity and reliability of the VRAG in a forensic psychiatric medium security population in Flanders. *Psychology, Crime & Law*, *22*(6), 530-537. doi: 10.1080/1068316X.2016.1168423

Jeandarme, I., Wittouck, C., Laenen, F. Vander, Ampe, M., Grouwels, Y., De Varé, J., Oei, T. I., Groenhuijsen, M., & Bogaerts, S. (2015). Forensische pilootprojecten 'medium security': Incidenten tijdens behandeling van geïnterneerden ressorterend onder CBM Gent [Forensic 'medium security' projects: Registration of incidents during treatment of forensic psychiatric patients of the Commission of Social Defense Ghent]. *Panopticon*, *36*(1), 26-45.

Jeandarme, I., Pouls, C., Wittouck, C., Laenen, F. Vander, Ampe, M., Verelst, R., Degrauwe, S., Oei, T. I., & Bogaerts, S. (2015). Forensische pilootprojecten 'medium security': Herval in delictgedrag na behandeling van geïnterneerden ressorterend onder CBM Gent [Medium security units in Flanders: Relapse after treatment of forensic psychiatric patients under the authority of the Commission of the Protection of Society Ghent]. *Panopticon*, *36*(3), 227-247.

Pouls, C., & **Jeandarme, I.** (2015). Risk assessment and risk management in offenders with intellectual disabilities: Are we there yet? *Journal of Mental Health Research in Intellectual Disabilities*, *8*(3-4), 213-236. doi: 10.1080/19315864.2015.1070221

Uzieblo, K., Habets, P., & **Jeandarme, I.** (2015). De IQ-score in vraag gesteld: Evolutie naar een multidimensioneel cognitief vaardigheidsprofiel. Een illustratie uit de forensische psychiatrie. *Neuropraxis*, *19*(4), 87-94. doi: 10.1007/s12474-015-0095-y

Habets, P., **Jeandarme, I.**, Uzieblo, K., Oei, T. I., & Bogaerts, S. (2014). Intelligence is in the eye of the beholder: Investigating repeated IQ measurements in forensic psychiatry. *Journal of Applied Research in Intellectual Disabilities*, *28*(3), 182-192. doi: 10.1111/jar.12120

Pouls, C., & **Jeandarme, I.** (2014). Psychopathy in offenders with intellectual disabilities: A comparison of the PCL-R and PCL:SV. *International Journal of Forensic Mental Health*, *13*(3), 207-216. doi: 10.1080/14999013.2014.922138

Pouls, C., **Jeandarme, I.**, & Habets, P. (2014). Risicotaxatie bij daders met een verstandelijke beperking. Eerste toepassing Nederlandstalige VRAG [Risk assessment in offenders with intellectual disabilities: A first application of the Dutch VRAG]. *De Psycholoog*, *1*, 42-51.

Stassen, W., Habets, P., Mertens, A., Laender, J. De, & **Jeandarme, I.** (2014). The InReach project: From penitentiary to forensic hospital. *The International Journal of Therapeutic Communities*, *35*(3),119-126. doi: 10.1108/TC-01-2014-0002

Jeandarme, I. (2002). Derde generatie antidepressiva. *Psychiatrie en verpleging*, *78*(2), 83-98.

BOOK CHAPTERS

Goethals, K., Dhoore, T., Groot, A. de, **Jeandarme, I.**, Keulen-de Vos, M., Pouls, C., Soe-agnie, S., & Willemsen, J. (2015). Differentiële diagnostiek en comorbiditeit bij psychopathie en de antisociale persoonlijkheidsstoornis. In W. Canton, L. Claes, L. Gijs, I. Jeandarme, E. Klein-Haneveld, & D. van Beek (Eds.), *Handboek antisociale persoonlijkheidsstoornis en psychopathie* (pp. 315-358). Utrecht: De Tijdstroom.

Peters, M. J. V., Hendrickx, J., **Jeandarme, I.**, & Pouls, C. (2012). Psychopathy and risk assessment in intellectually disabled forensic patients: old problems and new insights. In K. Oei & M. Groenhuijsen (Eds.), *Progression in Forensic Psychiatry: About Boundaries* (pp. 377-396). Deventer: Kluwer.

Peters, M., **Jeandarme, I.**, & Pouls, C. (2012). Recidive en risicotaxatie binnen de forensisch psychiatrische zorg: Een huidige stand van zaken en nieuwe ontwikkelingen. In. J. Mostinckx, F. Deven, & P. Dauwe (Eds.), *Welzijn en Zorg in Vlaanderen: Wegwijs voor de sociale sector 2012-2013* (pp. 471-486). Mechelen: Kluwer.

Jeandarme, I. (2010). Risicotaxatie bij psychiatrische patiënten: Risky business. In G. Benoit, J. De Fruyt, H. Nys, G. Rommel, G. Steegen, P. Van Peteghem, & J. Van Speybroeck (Eds.), *De bescherming van de persoon van de geesteszieke* (pp. 141-158). Brugge: die Keure.

Putte, D. Van de, & **Jeandarme, I.** (2009). Agressief gedrag. In M. W. Hengeveld, A. J. L. M. van Balkom, C. van Heeringen, & B. G. C. Sabbe (Eds.), *Leerboek Psychiatrie* (pp. 503-512). Utrecht: De Tijdstroom.

Ridder, H. De, **Jeandarme, I.,** & Timmerman, K. (1994). Druggebruik in de gevangenis: problemen en hulpverlening. In Christiaensen, S. & Goethals, J. (Eds.), De illegale drugsgebruiker tussen strafrechtspleging en hulpverlening (pp. 299-318). Leuven: Acco.

BOOKS

Canton, W., Claes, L., Gijs, L., **Jeandarme, I.**, Klein-Haneveld, E., & Beek, D. van (2015). *Handboek antisociale persoonlijkheidsstoornis en psychopathie*. Utrecht: De Tijdstroom.

Jeandarme, I., Pouls, C., & Peters, M. J. V. (2012). *Violence Risk Appraisal Guide: Richtlijnen om het risico op (seksueel) gewelddadig gedrag te beoordelen*. Hasselt: Leën.

OTHER PUBLICATIONS

Jeandarme, I. (2015). The Virginia Tech Massacre. Strategies and challenges for improving mental health policy on campus and beyond [Bespreking van het boek *The Virginia Tech Massacre. Strategies and challenges for improving mental health policy on campus and beyond*, door A. B. Sood & R. Cohen]. *Tijdschrift voor Psychiatrie*, *57*(8), 622.

Jeandarme, I. (2015). Forensische psychiatrie is ook psychiatrie. *Periodiek*, *70*(1), 24-25.

Jeandarme, I., Wittouck, C., Laenen, F. Vander, Varé, J. De, Grouwels, Y., Ampe, M., & Oei, T. I. (2013). Disciplinary infractions in three Flemish forensic institutions: Prevalence and judicial reaction. In P. Callaghan, N. Oud, J. H BjØrngaard, H. Nijman, T. Palmstierna, R. Almvik, & B. Thomas (Eds.), *Proceedings of the 8th European Congress on Violence in Clinical Psychiatry* (pp. 391-394). Amsterdam: Kavanah.

Pouls, C., & **Jeandarme, I.** (2013). Institutional aggression in intellectually disabled (ID) offenders – An analysis of 5 specialized forensic ID projects. In P. Callaghan, N. Oud, J. H BjØrngaard, H. Nijman, T. Palmstierna, R. Almvik, & B. Thomas (Eds.), *Proceedings of the 8th European Congress on Violence in Clinical Psychiatry* (pp. 408-411). Amsterdam: Kavanah.

Jeandarme, I. (2012). Komen mannen echt van Mars en vrouwen van Venus? [Are men really from Mars and women from Venus?]. *Tijdschrift voor Psychiatrie*, *54*(4), 339-340.

Jeandarme, I. (2008). PCSOT: De urinetest bij de behandeling van zedendelinquenten? *Tijdschrift voor Seksuologie*, *32*, 217-218.

SUBMITTED ARTICLES

Jeandarme, I., Wittouck, C., Laenen, F. Vander, Pouls, C., Heimans, H., Oei, T. I., & Bogaerts, S. (accepted pending minor revisions). Critical incidents and judicial response during medium security treatment. *International Journal of Law and Psychiatry.*

Jeandarme, I., Pouls, C., Oei, T. I., & Bogaerts, S. (accepted pending minor revisions). Forensic psychiatric patients with comorbid psychopathy: Double trouble? *International Journal of Forensic Mental Health*.

Jeandarme, I., & Schipaanboord, E. (2016). Internering. In K. Goethals, G. Meynen, & A. Popma (Eds.), *Leerboek forensische psychiatrie*. Manuscript in review.

Uzieblo, K., Pouls, C., Habets, P., & **Jeandarme, I.** (2016). Plegers met een intellectuele disfunctie. In K. Goethals, G. Meynen, & A. Popma (Eds.), *Leerboek forensische psychiatrie.* Manuscript in review.

Smid, W., Berg, J. W. van den, & **Jeandarme, I.** (2017). Plegers van seksueel geweld. In L. Gijs, L. Aerts, P. Enzlin, M. Dewitte, J. Georgiadis, B. Kreukels en E. Meuleman (Eds.), *Handboek seksuologie*. Manuscript in review.

Jeandarme, I., Pouls, C., Al-Taiar, H., Brink, J., Canton, W., Kristiansson, M., . . . Konrad, N. (2016.) *Criminal responsibility evaluations: Benchmarking in different countries*. Manuscript submitted for publication.

Hornsveld, R., Kraaimaat, F., Gijs, L., Zwets, A., **Jeandarme, I.,** & van Marle, H. (2016). *Behandelprogramma voor seksueel geweldplegers in forensisch psychiatrische instellingen: Voorstel voor een onderzoek in Nederland en Vlaanderen*. Manuscript submitted for publication.

Jeandarme, I., Vandenbosch, L., Groenhuijsen, M., Oei, T.I., & Bogaerts, S. (2016). *Who are the victims of NGRI acquittees? A study among Belgian internees*. Manuscript submitted for publication.